Journal of Communication Disorders 122 (2026) 106657

Contents lists available at ScienceDirect

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis

A multilevel linguistic analysis reveals linguistic impairments in persons with multiple sclerosis

Marika Gobbo [a,b], Ilaria Gabbatore [c,g], Giulia Piovani [d], Francesca Marina Bosco [e,g], Andrea Marini [a,f,1,*], Andrea Tacchino [d,1]

[a] *Cognitive neuroscience Lab, University of Udine, Udine, Italy*
[b] *Department of Life Sciences, University of Trieste, Trieste, Italy*
[c] *Department of Humanities, University of Turin, Torino, Italy*
[d] *NeuroBRITE Research Center, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy*
[e] *Department of Psychology, University of Turin, Turin, Italy*
[f] *Scientific Institute, IRCCS E. Medea, Pasian di Prato (UD), Italy*
[g] *GIPSI - research Group on Inferential Processes in Social Interaction, University of Turin, Turin, Italy*

ARTICLE INFO

Keywords:
Multiple sclerosis
Multilevel linguistic analysis
Language impairments
Discourse production

ABSTRACT

Background: Multiple sclerosis results in sensorimotor and cognitive deficits. Previous studies have provided limited evidence of linguistic deficits, largely due to a lack of comprehensive assessments. This study addressed this gap by applying a Multilevel procedure of Linguistic Analysis (MLA) of discourse production in persons with MS (PwMS).
Methods: 20 PwMS and 20 healthy controls, matched for age, education and sex, were administered tasks assessing cognitive skills and narrative discourse production abilities. The speech samples were analyzed using the MLA to assess both micro- and macrolinguistic skills.
Results: The MLA highlighted significant linguistic deficits in PwMS, including frequent phonological errors, reduced grammatical accuracy, limited lexical informativeness, and impaired discourse cohesion and coherence. Microlinguistic difficulties were found to correlate with macrolinguistic impairments, confirming their interconnected nature. Fatigue level significantly affected performance in PwMS. Interestingly, no significant correlations emerged between linguistic and cognitive measures.
Conclusions: The comprehensive assessment of the narrative abilities of PwMS showed their impairments span lexical, grammatical, and discourse-level processes. These findings are discussed considering recent neuroimaging evidence about the neural underpinnings of discourse production and underscore the importance of incorporating training programs that target both micro- and macrolinguistic components of narrative production into clinical care's paths.

1. Introduction

Multiple sclerosis (MS) is an autoimmune disease that damages the myelin of the axons of the central nervous system (Renauld et al., 2016). The primary clinical courses of MS are Relapsing-Remitting (RR), Secondary-Progressive (SP), and Primary-Progressive

* Corresponding author at: Cognitive Neuroscience Laboratory, University of Udine, Via Margreth, 3, 33100 Udine, Italy.
E-mail address: andrea.marini@uniud.it (A. Marini).
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.jcomdis.2026.106657
Received 29 May 2025; Received in revised form 2 March 2026; Accepted 7 May 2026
Available online 12 May 2026
0021-9924/

(PP). These phenotypes are defined based on the presence or absence of clinical activity, including relapses and disease progression, as well as magnetic resonance imaging (MRI) findings, such as new lesions indicating inflammatory activity and atrophy suggesting ongoing neurodegeneration (Aj et al., 2018). The most common symptoms include physical and sensorimotor issues (e.g., fatigue, sensory disturbances, spasticity, and balance problems), as well as cognitive impairments (Portaccio & Amato, 2022), particularly those affecting attention (Arnett & Strober, 2011; Beatty et al., 1996; Litvan et al., 1988), information processing (Diamond et al., 1997; Litvan et al., 1988; Migliore et al., 2018), executive functions (Arnett & Strober, 2011; Denney et al., 2005; Foong et al., 1997), processing speed (Chiaravalloti & DeLuca, 2008; DeLuca et al., 2004), long-term memory (Rao, 1995), and speech production (Plotas et al., 2023).

1.1. Linguistic impairments in MS

To date, only few studies have expanded their focus to include the assessment of linguistic impairments in persons with MS (PwMS) (e.g., Arnott et al., 1997; Carotenuto et al., 2018; Grigoriadis et al., 2024; Rao, 1995; Renauld et al., 2016). A comprehensive linguistic assessment should account for the fact that language is a complex cognitive system, requiring the integration of both micro- and macrolinguistic levels of processing (Marini, Andreetta et al., 2011). The microlinguistic level pertains to the processing of lexical and grammatical information, while the macrolinguistic level involves contextualizing linguistic information through pragmatic processing and organizing propositions into cohesive and coherent discourse. With respect to language production, growing evidence suggests that PwMS may exhibit weaknesses at both micro- and macrolinguistic levels.

At the microlinguistic level, difficulties have been reported specifically in morphosyntactic processing (Fyndanis et al., 2020; Grigoriadis et al., 2024). Notably, a study by Grigoriadis and colleagues (2024), involved administering two tasks to PwMS: one assessing morphological production (e.g., sentences requiring a verb in the past tense) and another evaluating syntactic comprehension (e.g., pointing to the picture corresponding to the meaning of a given direct object, relative, or passive sentence). People in the clinical group scored significantly lower than Healthy Controls (HC) in morphological production for both regular and irregular verbs. In contrast, differences in syntactic comprehension between the two groups were observed only for passive sentences (as opposed to subject, object, or reflexive sentences), with PwMS scoring lower than the HC.

At the macrolinguistic level, where pragmatics (Levinson, 1983) is conceptualized as a higher-order cognitive-linguistic domain (Bara, 2010) rather than a purely functional discourse outcome (see Dutta et al., 2024), pragmatic impairments in PwMS were thoroughly described by Carotenuto and colleagues (2018). The authors administered an Italian standardized test for pragmatic abilities (i.e., Assessment of Pragmatic Abilities and Cognitive Substrates - APACS (Arcara & Bambini, 2016)) to 42 PwMS and 42 HC. Based on the overall total score, 55 % of participants in the clinical group exhibited pragmatic impairments that affected their production and comprehension of non-literal language. Notably, in the Interview and Narrative subtasks of the APACS, PwMS performed worse than HC. Regarding discourse production abilities in PwMS, Arnott and colleagues (1997) assessed the narrative skills of 47 PwMS by asking them to describe a sequence of cartoons. The narratives they produced differed significantly from those produced of HC in terms of core propositions (i.e., propositions produced by at least 20% of the control group). As a result, PwMS provided insufficient essential information when describing the stories. Similarly, in a study by Arrondo and colleagues (2009), 16 PwMS were asked to produce a 30 min monologue about their life history. People in the clinical group produced fewer words, organized in shorter sentences. Interestingly, the evaluator had to interact more frequently with the PwMS to sustain the conversation.

Taken together, these studies suggest that PwMS may experience linguistic difficulties that hinder effective and informative communication. Interestingly, these findings align with PwMS' own perceptions about their linguistic impairments. It is worth noting that such awareness is not always present in individuals with linguistic impairments. For instance, in clinical conditions like Wernicke's aphasia, there is often limited awareness of language and communication difficulties. In a study by El-Wahsh and colleagues (2020), the *Speech pathology-specific questionnaire for persons with Multiple Sclerosis* (SMS; (El-Wahsh et al., 2019)) was used to examine the prevalence and nature of self-reported linguistic impairments in PwMS. Seventy-five percent of the 160 participants who completed the questionnaire self-reported experiencing linguistic difficulties. Notably, the most common symptoms involved challenges with word retrieval and spoken discourse. Furthermore, the linguistic difficulties of PwMS may also be related to impairments in cognitive functions (e.g., working memory, sustained attention, and executive functions) that – also in other populations - have been found to significantly contribute to narrative discourse production (e.g., for a general review on this topic, see Marini et al., 2023; for persons with traumatic brain injuries, Büttner-Kunert et al., 2022; for persons with aphasia, Dutta et al., 2024; for healthy older adults: Hilviu et al., 2025). Working memory enables the temporary storage and manipulation of information and comprises multiple components (Baddeley, 2012). Sustained attention is the ability to maintain focus over time. Importantly, both working memory and sustained attention are crucial for the efficiency of the set of higher-order cognitive skills labeled under the umbrella term of Executive Functions (EFs). These include the abilities to plan actions, monitor performance via sustained attention, update incoming information through working memory, and exert inhibitory control (Miyake et al., 2000). Difficulties in working memory, sustained attention, and executive functions may contribute to reduced discourse efficiency as it has been shown, for example, in healthy aging (Merlini et al., 2026) as well as in persons with schizophrenia (Marini et al., 2008) and traumatic brain injuries (Marini et al., 2017).

1.2. The neural correlates of linguistic impairments in MS

Linguistic difficulties in these persons may be linked to the neurofunctional alterations caused by MS. Although the neural correlates of discourse production are still under investigation, it is likely that this process involves a broad frontal-parietal network. This network spans from classical left perisylvian areas to language-related structures that regulate domain-general cognitive functions (e.

g., attention, executive functions) (Braun et al., 2001; Spalletta et al., 2010). Within this network, evidence suggests that the left Inferior Frontal Gyrus (lIFG) may play a key role in generating and producing informative words in narrative discourse (Marini & Urgesi, 2012; Mazzon et al., 2019). In PwMS, both functional and structural alterations in these networks have been documented. For example, left frontal lesions have been associated with difficulties in abstract problem-solving, working memory, and lexical fluency, while lesions in the left parieto-occipital cortex have been linked to deficits in verbal learning and complex visual-integrative skills (Swirsky-Sacchetti et al., 1992). Moreover, a recent review by Rocca and colleagues (2022) demonstrated that in PwMS the absence of cognitive impairments is associated with increased frontal-parietal activity. However, when cognitive deficits are present, these people exhibit reduced activity in the same regions. Given that this is a broad network encompassing the dorsolateral prefrontal, premotor, supplementary motor, medial, and parietal cortices, the pathological mechanisms of MS might disrupt brain connectivity, leading to linguistic impairments.

1.3. Limitations of the available evidence on language impairments in MS

These findings suggest that linguistic difficulties experienced and reported by PwMS should not be underestimated in clinical practice. Addressing language impairments and planning effective interventions, including speech and language therapy (SLT), necessitates an accurate assessment and a detailed understanding of strengths and weaknesses. In this context, it is notable that, of the 120 PwMS who reported a linguistic impairment in the study by El-Wahsh and colleagues (2020), only one was receiving SLT at the time of the study. Therefore, a specific assessment of narrative abilities in PwMS could enhance understanding of the micro- and macro-linguistic skills affected in this population and support the development of targeted rehabilitative programs.

Discourse production is fundamental to everyday communication, as it supports activities such as participating in conversations, narrating personal experiences, giving instructions, and engaging with others, all of which require the integration of linguistic, cognitive, and social-pragmatic processes (Bryant et al., 2016; Kong et al., 2025; Labov & Waletsky, 1967; Marini et al., 2022). Given its importance in daily life activities, further research on this aspect of complex language processing is urgently needed. To the best of our knowledge, no study has yet provided a detailed characterization of narrative impairments affecting both micro- and macrolinguistic levels of processing in MS. Moreover, considering the self-reported challenges that PwMS face in expressing themselves through spoken discourse, such investigations are essential to promote their inclusion in tailored rehabilitation programs focused on language and communication. Indeed, language-related symptoms in MS can have negative psychosocial ramifications, such as frustration, embarrassment, loneliness, limitations in social and familial relationships, and, consequently, can reduce mental and physical health-related quality of life (El-Wahsh et al., 2020, 2021; Kristensson et al., 2022; Yorkston et al., 2007).

Previous studies (Arnott et al., 1997; Arrondo et al., 2009) investigating narrative abilities in PwMS have primarily focused on measures of informativeness (e.g., core propositions), word counts, and mean sentence length, without examining all aspects of micro- and macrolinguistic processing. However, efficient discourse production relies on the continuous interaction between these two levels. Increasing evidence suggests that such interactions are evident in both healthy individuals (e.g., Marini et al., 2025; Merlini et al., 2026) and persons with different etiologies (e.g., for persons with anomic aphasia, Andreetta et al., 2012; for persons with schizophrenia, Marini et al., 2008; for persons with right hemisphere damage, Marini, 2012). For example, Andreetta et al. (2012) showed that persons with aphasia may experience macrolinguistic difficulties (e.g., in the generation of cohesive ties among utterances) because of underlying lexical difficulties (e.g., anomias that may induce abrupt interruptions of ongoing utterances and determining also a reduction in the grammatical complexity of their utterances). Similarly, persons with schizophrenia have been shown to experience significant difficulties in inhibitory control that determined the inclusion of several tangential utterances in their narratives. This reduced their levels of informativeness but also determined the frequent production of semantic and even morphological errors in their utterances (Marini et al., 2008). In this context, it would be highly valuable to explore whether microlinguistic impairments in PwMS influence the macrolinguistic aspects of discourse production.

1.4. Aims of the current study

Given the limitations outlined above, the present study aimed to provide a detailed analysis of micro- and macrolinguistic abilities in PwMS by comparing their narrative discourse to that of a group of HC. Specifically, we applied a Multilevel procedure of Linguistic Analysis (MLA; Marini, Andreetta et al., 2011, see Marini et al., 2026 for the recent standardization) to language samples collected during a picture-description task. Previous studies suggest that this method is effective in detecting linguistic impairments across various conditions, including acute and chronic post-stroke aphasia (Andreetta et al., 2024; Marini, Andreetta et al., 2011), traumatic brain injury (Marini, Galetto et al., 2011), Autism Spectrum Disorders (Adornetti et al., 2023), schizophrenia (Marini et al., 2008), Alzheimer disease (Mazzon et al., 2019), and during inhibitory non-invasive brain stimulation in healthy adults (Marini & Urgesi, 2012). Therefore, the MLA may serve as a valuable tool for identifying and characterizing linguistic impairments in PwMS.

The study was designed to address the following research questions: (1) Is it possible to identify impaired linguistic skills in PwMS using the MLA of discourse production? (2) Can we determine whether micro- and macrolinguistic difficulties are interrelated in these individuals? (3) Are there potential relationships between altered linguistic abilities and impaired cognitive functions in PwMS? (4) What impact does self-perceived fatigue level exert on cognitive and narrative production skills in PwMS? In the present study, self-perceived fatigue was conceptualized as a clinically meaningful explanatory variable indexing intra-group heterogeneity within PwMS, rather than as a confounding factor to be statistically controlled. Given the well-documented impact of fatigue on cognitive efficiency in MS, we aimed to determine whether distinct fatigue profiles would be associated with differential cognitive and narrative performance. We hypothesized that (i) the MLA would allow us to identify the specific linguistic levels affected by MS; (ii) the detailed nature of the MLA would allow us to capture the interplay between micro- and macrolinguistic difficulties; (iii) We would detect relationships between cognitive measures and linguistic impairments; and (iv) Participants with higher scores on a questionnaire assessing fatigue levels would have the lowest performance on tasks assessing cognitive and narrative production skills.

2. Materials and methods

2.1. Participants

Forty Italian-speaking adults participated in the study, consisting of 20 PwMS and 20 HC. The two groups did not differ significantly in age (t(38)= 0.117; $p = .713$), level of education ($U = 218.50$; p=.608; rank biserial correlation: 0.093), or sex distribution (X^2 (1, $N = 40$)= 0.100, p= .752, Cramérs $V = 0.050$). Socio-demographic variables (age, sex, years of education) and clinical data (disease duration, MS type and characteristics) were collected and are summarized in Table 1. Disability was assessed using the Expanded Disability Status Scale (EDSS; Kurtzke, 1983), a clinician-rated scale ranging from 0 to 10, with higher scores indicating greater neurological disability. EDSS scores ≤3.5, in the range 4.0–6.0, and ≥6.5 reflect mild, moderate and severe disability, respectively (Kobelt et al., 2006). Fatigue was measured using the Modified Fatigue Impact Scale (MFIS; Fisk et al., 1994), a 21-item self-reported questionnaire (score range 0–84) providing an assessment of the effects of fatigue in terms of physical, cognitive, and psychosocial functioning. Higher scores indicate greater fatigue impact. The sample included individuals with multiple sclerosis presenting moderate to severe neurological disability, with all participants but one showing EDSS scores ≥4. Despite this homogeneous level of disability, perceived fatigue was highly variable, with MFIS total scores ranging from 9 to 64, indicating the presence of distinct fatigue profiles within the sample. A significant difference was found between PwMS and healthy controls in total MFIS score (t(34)= -5.48; $p < .001$).

Given the variability in fatigue impact, participants were further stratified into two subgroups based on MFIS total scores, reflecting lower (PwMS_Low; MFIS < 38) versus higher fatigue impact (PwMS_High; MFIS ≥ 38) (see Larson, 2013; Table 1). The two subgroups had comparable age (t(18)= 0.992; p= .334, $d = 0.453$), level of education (t(18)= 1.345; p= .195, $d = 0.614$), sex distribution [X^2 (1, $N = 20$)= 0.135, p= .714, Cramérs $V = 0.082$], disease duration (t(18)= 1.256; p= .225, $d = 0.573$), and EDSS (t(18)= 0.188; p= .853, $d = 0.086$).

PwMS were recruited through the NeuroBRITE Research Center (Italian Multiple Sclerosis Foundation, FISM) at the AISM Rehabilitation Service of Liguria (Genoa, Italy). Participants in the control group were recruited on a voluntary basis via various methods, including local associations, cultural events, personal contacts, and advertisements posted on the primary social media platforms (e.g., Facebook) of the research groups involved.

Inclusion criteria were as follows: (a) participants had to be at least 18 years old; (b) native Italian speakers; (c) possess basic cognitive and linguistic abilities to ensure that any observed “decline” at the narrative level was not attributable to foundational deficits. PwMS were included if they achieved the cut-off scores on the following neuropsychological tests: Mini Mental State Examination (MMSE; (Folstein et al., 1975); cut-off score: ≥ 24/30), Aachener Aphasie Test – naming subtest (AAT, (Luzzatti et al., 1996), 1983; cut-off score: 108/120), and Token test ((De Renzi & Vignolo, 1962); cut-off score ≥ 4.5/6). Additional criteria included: (d) absence of concomitant neurological or psychiatric diagnoses; (e) no prior history of neurological or psychiatric disorders; and (f) no drug or alcohol addiction. Specific inclusion criteria for PwMS were: (g) a confirmed diagnosis of MS and (h) an EDSS score between 1 and 6.5. All participants were informed about the study’s aims and procedures and provided informed consent. The two groups were matched on age, gender, and level of formal education (see Table 1). The study was approved by the ethical committee of A.O. Universitaria San Martino of Genoa, Italy.

Table 1
Demographic and clinical characteristics of the participants. Results are presented as means, (standard deviations), and ranges for the groups of healthy controls and PwMS and, separately, for the two subgroups of PwMS with high and low MFSI total scores. The asterisk (*) shows when the group-related difference was significant. For disease duration the years since the first diagnosis are reported. *Legend:* PwMS= Persons with Multiple Sclerosis; PwMS_High= Persons with Multiple Sclerosis and a high level of fatigue; PwMS_Low= Persons with Multiple Sclerosis and a low level of fatigue; HC= Healthy Control participants; PP= primary progressive; SP= secondary progressive; RR= relapsing remitting; EDSS= Expanded Disability Status Scale; MFIS= Modified Fatigue Impact Scale.

	PwMS	PwMS_High	PwMS_Low	HC
N°	20	12	8	20
Sex (female/male)	11/9	7/5	4/4	10/10
Age (y)	57.65 (9.1); range: 35–78	57.58 (7.99); range 47–78	57.75(11.23); range 35–70	56.55 (9.60); range 38–75
Education (y)	12.05 (4.0); range 8–18	12.60 (3.84); range 8–18	11.08 (3.87); range 8–18	12.6 (3.80); range 8–18
Disease duration (y)	23.00 (11.50); range 7–43	21.00 (9.75); range 7–38	27.50(13.46); range 8–43	
Subtype	PP:3; SP:12; RR:5	Pp:1; SP:8; RR:3	Pp:2; SP:4; RR:2	
EDSS	5.5 (1.2); range 2–6.5	5.54 (0.96); range 4–6.5	5.44 (1.52); range 2–6.5	
MFIS total*	39.95 (17.1); range 9–64	51.58 (9.27); range 38–64	22.50 (8.85); range 9–33	12.19 (12.1); range 0–44

2.2. Neuropsychological assessment

To ensure the inclusion of PwMS who have adequate baseline cognitive levels and without linguistic comprehension or production issues on traditional tests, three baseline assessments were administered to evaluate overall cognitive function, naming skills, and language comprehension. Specifically, overall cognitive impairment was assessed using the Mini-Mental State Examination (MMSE; (Folstein et al., 1975), a screening tool that evaluates orientation, registration, attention, calculation, recall, language, and visuospatial abilities. Lexical production difficulties were evaluated using the naming subtest of the Italian version of the Aachener Aphasie Test (AAT; (Luzzatti et al., 1996). Linguistic comprehension difficulties were assessed with a shortened version of the Token Test (De Renzi & Vignolo, 1962). In addition to the baseline evaluation, further tests were administered to assess working memory, attention, and executive functions using tasks that predominantly rely on verbal and linguistically mediated stimuli. Working memory and visual attention were assessed using Digit recall tasks and the Trail Making Test, respectively, as these measures have been shown to tap executive-working memory efficiency. In a recent study of >300 healthy adults, performance on Forward and Backward Digit Recall, TMT_A, and the TMT_B-A index loaded onto a common executive-working memory efficiency component, which was negatively associated with narrative measures reflecting semantic coherence breakdown, utterance formulation difficulties, and reduced productivity (Merlini et al., 2026).

2.2.1. Working memory

Working memory was assessed using the Forward and Backward Digit Recall tests (David, 1987; Monaco et al., 2013). In the Forward Digit Span task, participants were required to repeat a series of numbers read aloud by the examiner. The sequences began with three digits and progressively increased in length to a maximum of nine digits, with each sequence presented twice. Participants earned one point for each correctly repeated sequence, and their score corresponded to the longest sequence they successfully repeated at least once. In the Backward Digit Span task, participants were instructed to repeat the numbers in reverse order. The sequence length started at three digits and increases to eight digits. As in the forward task, one point was awarded for each correct attempt, with the score reflecting the longest reverse sequence repeated correctly at least once.

2.2.2. Attention and executive functions

The Trail Making Tests (TMT; Giovagnoli et al., 1996) evaluates sustained visual attention (in both Part A and Part B, referred to as TMT-A and TMT-B, respectively) and set-shifting and cognitive flexibility (TMT-B). In TMT-A, participants are required to connect a set of randomly positioned numbered circles in numeric order (from 1 to 25) as quickly as possible. In TMT-B, participants must alternatively connect a set of randomly positioned circles (1 to 13) and letters (A to N) in numeric-alphabetic order (e.g., 1-A-2-B) as quickly as possible. The score for each part is determined by the time taken to complete it. Additionally, the difference in completion time between TMT-B and TMT-A (i.e., TMT_B-A) is often used as an indicator of inhibitory control (Arbuthnott & Frank, 2000).

The Wisconsin Card Sorting Test (WCST; Nelson, 1976) was used to assess executive functions, including abstract reasoning, problem-solving, and cognitive flexibility. The test consists of 4 'stimulus cards' and two sets of 24 'response cards'. Each card displays a varying number (one to four) of geometric shapes (crosses, circles, stars or triangles) in different colors (green, blue, red or yellow). Each response card shares one characteristic (shape, color, or number) with three of the stimulus cards, while it has no characteristic in common with the fourth.

Participants were instructed to place the response cards under the stimulus cards, according to a specific rule, which was not disclosed to them. They were required to discover the rule through trial and error, with feedback from the examiner indicating whether their responses were correct.

2.3. Narrative production assessment

Narrative production was assessed using language samples obtained with a picture description task.

This task was selected because it provides a controlled yet naturalistic context for eliciting connected speech, allowing for systematic comparison across participants while minimizing variability due to differences in personal experience, memory, or emotional salience that may affect more open-ended elicitation methods (e.g., autobiographical narratives or conversations). Specifically, participants were asked to describe five pictures: three single scenes, namely the "Picnic" from the Western Aphasia Battery (Kertesz, 1982), the "Cookie Theft" by Goodglass and Kaplan (1972), the "Smith's family" by Schuell et al. (1964), and two sequences, the "Flowerpot" by Huber and Gleber (1982), and the "Quarrel" by Nicholas and Brookshire (1993). These specific pictures were chosen because they have been extensively used in previous studies to assess discourse production skills in healthy individuals and persons with brain disorders (e.g., Boucher et al., 2022; Brookshire & Nicholas, 1994; Marini et al., 2025; Merlini et al., 2026; Marcotte et al., 2024). Administration and transcription procedures followed the methodology described in Marini et al. (2011). Speech samples were audio-recorded and transcribed verbatim, including phonological fillers, pauses, false starts, and extraneous utterances. The MLA was then applied to the transcriptions.

Utterances were segmented based on four main criteria outlined in Marini et al. (2011): acoustic, semantic, grammatical, and phonological. The acoustic criterion defines an utterance as a segment of speech interrupted by a long pause (e.g., "*There is a … (10 sec) / kite*"). The semantic criterion considers an utterance when the proposition is complete and not interrupted (e.g., "*There is a kite in the sky*"). The grammatical criterion considers an utterance when it is grammatically complete and there are no clear pauses or propositional violations, it can also include subordinate clauses (e.g., "*There is a kite in the sky while a couple is having a picnic*"). The phonological criterion considers an utterance when a false start interrupts a word and consequently the completing of the sentence (e.

g., “*There is a ka-/ki-/kite in the sky*”).

The duration of the speech sample obtained in each story was also calculated. The transcripts of each story were then analyzed with respect to productivity, lexical and grammatical processing, and discourse organization. *Productivity* was calculated in terms of number of words, speech rate, and Mean Length of Utterance (MLU). *Lexical processing* was calculated in terms of percentage of phonological, semantic, and bound paragrammatic errors. *Grammatical processing* was calculated in terms of percentage of free paragrammatic errors, omissions of function words, omissions of content words, and grammatical completeness. *Discourse organization* was calculated in terms of percentage of lexical informativeness, cohesion errors, and local and global coherence errors (see Fig. 1 and Table 2).

2.4. Procedure

All participants were tested individually. PwMS were assessed at the AISM Rehabilitation Service of Liguria in Genoa, while HC were tested at the University of Turin. The assessment consisted of three 1 h sessions, scheduled based on participant availability. This structure was designed to minimize the impact of fatigue on performance, with participants encouraged to take breaks whenever needed.

The administration and scoring procedures were conducted by a research assistant who had been trained on the tools and protocols. Sessions were video recorded to facilitate offline coding by independent external raters. Both administration and scoring procedures followed the instructions provided in the manual for each task to ensure objective and standardized measurement. All scores were subsequently compiled into a dataset for statistical analysis.

Scoring was carried out by two expert raters (M.G. and A.M.). Both raters were blinded to the participants’ groups and preliminarily analyzed 35 narratives. Interrater reliability was assessed using the Kappa statistic (Carletta, 1996). Agreement between the two raters was almost perfect across all measures and substantial for % phonological errors (see Table 3).

2.5. Statistical analyses

To determine the presence of differences between participants with MS and healthy controls in pretests, cognitive tests, and narrative production measures a series of independent *t*-tests or Mann-Whitney U tests were conducted where appropriate. Specifically, the assumptions of normality and homogeneity of variance were evaluated using the Shapiro-Wilk test and the Levene's Test of Equality of Variances, respectively. When normality was violated, the non-parametric Mann-Whitney U test was used as an alternative to *t*-tests, while Welch’s adjusted t-statistic was applied in cases of variance violations. In these analyses, effect sizes were calculated in terms of Cohen’s d for *t*-tests (0.01 small effect; 0.06 medium effect; 0.14 large effect) and rank biserial correlations (r_{bs}) for *U tests* (0.10 small effect; 0.30 medium effect; 0.50 large effect). Alpha value was corrected using Benjamini-Hochberg adjustment for multiple comparisons.

To further assess the impact of fatigue on their performance and given the unequal group sizes (HC, PwMS_High and PwMW_Low), fatigue was conceptualized as a clinically meaningful explanatory variable reflecting intra-group heterogeneity within PwMS rather than as a confounding factor to be statistically controlled. Given the modest sample size and unequal group sizes, non-parametric group-based analyses were considered more robust than regression-based modeling of fatigue as a continuous predictor. Therefore, a series of non-parametric Kruskal-Wallis tests were conducted with group as the between subject factor across pretest, cognitive, and narrative variables. Effect sizes were calculated using rank ε^2, which estimates the proportion of variance in ranked data explained by group membership in Kruskal-Wallis tests (0.01 small effect; 0.06 medium effect; 0.14 large effect). When a significant group effect was observed, Dunn’s post hoc comparisons were performed.

Next, to explore the relationship between cognitive-linguistic functions and narrative abilities, a series of Pearson’s product-moment correlation analyses was conducted separately for HCs and PwMS. Finally, further Pearson’s correlation analyses were

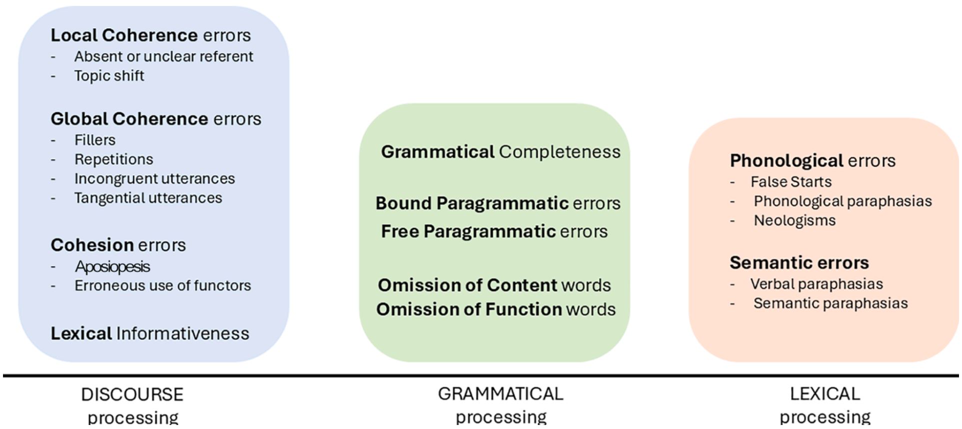

Fig. 1. Summary of measures of lexical, grammatical, and discourse processing obtained via the MLA.

Table 2
Schema showing the measures used in the MLA (adapted from Perlini et al., 2012).

MICROLINGUISTIC ANALYSIS	***Productivity***	Words	Total number of phonologically well-formed words, i.e., all those units of speech that were not scored as phonological errors
		Speech rate	Number of well-formed words produced in a minute (words/time)
		Mean length of utterance (MLU)	The mean number of words in each utterance (words/utterances)
	Lexical processing	% Phonological errors	Phonological errors included false starts (e.g., the attempt of the speaker to produce a target word that results in an intrasegmental interruption), phonological paraphasias (e.g., "pable" for "table"), and neologisms (i.e., an unrecognizable word due to multiple phonological errors). The percentage of phonological errors was calculated by dividing the number of phonological errors by the total number of units of speech and multiplying this value for 100.
		% Semantic errors	Semantic errors included semantic paraphasias (e.g., "chair" for "table"), and verbal paraphasias (e.g., "tomato" for "table"). The percentage of semantic errors was calculated by dividing the number of phonological errors by the total number of words and multiplying this value for 100.
		% Bound paragrammatic errors	Bound paragrammatic errors included the substitution of bound morphemes within a word (e.g., "*The girl is poured the drink*" where the bound paragrammatic error *poured* replaced the correct word *pouring*). The percentage of bound paragrammatic errors was calculated by dividing the number of these errors by the total number of words and multiplying this value for 100.
		% Free Paragrammatic errors	Free paragrammatic errors included the substitution of function words within a sentence. The percentage of free paragrammatic errors was calculated by dividing the number of these errors by the total number of words and multiplying this value for 100.
	Grammatical processing	% Omission Of function Words	Omnissions of function words were scored whenever a function word (e.g., articles, prepositions, pronouns) was missing in a sentence. The percentage of omissions of function words was calculated by dividing the number of omissions of function words by utterances and multiplying this value for 100.
		% Omission Of content Words	Omnissions of content words were scored whenever a content word (i.e., nouns, verbs, adjectives, or adverbs) was missing in a sentence. The percentage of omissions of content words was calculated by dividing the number of omissions of content words by utterances and multiplying this value for 100.
		% Grammatical completeness	A sentence was considered grammatically complete if all the elements of the argumental structure were present in the body of the sentence (i.e., without omissions of content words) and if there were no omissions in the morphosyntactic information (i.e., without omissions of function words) or substitution of function words or bound morphemes (i.e., without free paragrammatic errors and bound paragrammatic errors). The percentage of grammatical completeness was calculated by dividing the number of complete sentences by utterances and multiplying this value for 100.
MACROLINGUISTIC ANALYSIS		% Lexical informativeness	The percentage of words that were scored as lexical information units, i.e., those words that were not only well-formed from a phonological point of view, but also grammatically and pragmatically accurate. This broad category includes all those words that were not scored as phonological errors, semantic paraphasias, or paragrammatic errors, and were not ambiguous, repeated, or forming utterances scored as global coherence errors. Lexical informativeness was calculated as the percentage of words that were well-formed divided by number of words produced and multiplied for 100.
	Discourse organization	% Cohesion errors	Cohesion errors include the erroneous use of cohesive function words or semantically related content words and aposiopesis, i.e., the sudden interruption of an utterance with the completion of the grammatical information through the subsequent utterance (e.g., "*The children are trying to steal some … / They want to steal the cookies*)". More in detail, if the first utterance is incomplete because of the omission of morphosyntactic information, this omission causes the speaker to either reformulate the sentence in the subsequent utterance(s) or to omit pieces of information that may be important for the story. The percentage of cohesion errors was calculated by dividing the number of such errors by the number of utterances and multiplying this value for 100.
		% Local coherence errors	Local coherence errors comprise the use of words with absent or unclear referents and erratic topic shifts. A referent is unclear whenever the referent of a pronoun or the implicit subject of a verb is not unambiguously clear. A topic shift happens when an utterance is abruptly interrupted and the speaker instead of completing the preceding utterance, introduces a new topic (e.g., "*Children are trying to steal some … /The woman is washing the dishes*"). The percentage of local coherence errors was calculated by dividing the number of such errors by the number of utterances and multiplying this value for 100.
		% Global coherence errors	Global coherence errors refer to the production of utterances that might be tangential, conceptually incongruent with the story, propositional repetitions, or fillers. A tangential utterance provides information that is a derailment from the flow of discourse (e.g., *"Children are stealing some biscuits /I eat biscuits every*

(*continued on next page*)

Table 2 (*continued*)

	morning"). A conceptually incongruent utterance provides ideas that are not directly related to the stimulus (e.g., *"Children are stealing some biscuits /The woman is helping them").* A propositional repetition is a sequence of concepts and ideas that the speaker has already provided (e.g., *"Children are stealing some biscuits /They are stealing biscuits").* A filler utterance occurs when the speaker is not providing any additional information to the story, for example when the speaker verbalizes that he/she is thinking about something about the story or what to say (e.g., *"Children are stealing some biscuits /Is this correct?").* The percentage of Errors of Global Coherence was calculated by dividing the number of such errors by the number of utterances and multiplying this value by 100.

Table 3
Details of the Interrater Reliability Analysis performed on the narrative analysis conducted by two independent raters using the Kappa statistic.

Target variable	Rater 1	Rater 2	Kappa
Words	102.20 (70.54)	102.20 (70.54)	1.000
Speech rate	114.80 (24.25)	114.90 (24.20)	0.940
MLU	6.57 (1.54)	6.57 (1.54)	1.000
% Phonological errors	.89 (1.32)	1.11 (1.39)	0.786
% Semantic paraphasias	0.83 (2.12)	.83 (2.12)	1.000
% Bound morphological errors	.00 (0.00)	.00 (0.00)	–
% Free morphological errors	0.49 (0.853)	0.83 (2.28)	0.883
% Omissions content words	14.37 (11.49)	14.34 (11.51)	0.937
% Omissions function words	2.26 (4.76)	2.26 (4.76)	1.000
% Complete sentences	62.80 (17.71)	62.94 (17.76)	0.940
% Cohesion errors	21.09 (12.89)	21.06 (12.88)	0.969
% Local coherence errors	21.77 (15.25)	21.74 (15.21)	0.939
% Global coherence errors	11.14 (10.75)	11.14 (10.75)	1.000
% Lexical informativeness	82.00 (9.52)	81.83 (9.38)	0.909

conducted to investigate possible interactions between impaired macro- and microlinguistic measures in PwMS.

2.6. Sensitivity power analysis

Sensitivity power analyses were conducted using G*Power. The analysis revealed that with 20 participants per group, using a two-

Table 4
Cognitive and linguistic assessment of healthy controls and PwMS and, separately, for the two subgroups of PwMS with high and low MFSI total scores. Results are presented as means (standard deviations). The asterisks show when the group differences were significant after Benjamini–Hochberg correction for multiple comparisons. $p \leq .010$**; $p \leq .001$***. *Legend:* PwMS= Persons with Multiple Sclerosis; HC= Healthy Control participants; PwMS_High= Persons with Multiple Sclerosis and a high level of fatigue; PwMS_Low= Persons with Multiple Sclerosis and a low level of fatigue; WCST= Wisconsin Card Sorting Test; AAT= Aachener Aphasie Test; TMT= Trail Making Test.

Neuropsychological assessment		PwMS	PwMS_High	PwMS_Low	HC
MMSE		28.16 (2.10); range: 24–30	27.31 (2.15); range: 24–30	29.43 (1.27); range: 26–30	28.88 (1.45); range: 26–30
Naming	Naming subtest AAT	117.0 (2.27); range: 112–120	116.60 (2.15); range: 112–120	117.60 (2.45); range: 113–20	116.5 (3.20); range: 109–20
Comprehension	Token-Test	4.85 (0.33); range: 4–5	4.75 (0.40); range 4–5	5.00 (0.00); range 5–5	6.38 (5.83); range: 4–5
Working memory	Forward Digit Recall	5.83 (2.25); range 3–10	4.89 (1.91); range 3–8	7.29 (1.98); range 4–10	7.29 (2.01); range: 5–12
	Backward Digit Recall	5.99 (2.00); range 3–11	5.33 (1.61); range 3–9	6.99 (2.20); range 4–11	5.94 (1.83); range: 2–9
Attention	TMT-A (s)***	36.59 (18.15); range 18–97	39.45 (23.03); range 18–97	32.30 (5.11); range 25–41	23.46 (7.52); range: 10–38
	TMT-B (s)	72.75 (47.42); range 11–214	89.06 (53.19); range 18–214	48.29 (22.78); range 11–87	45.75 (28.57); range: −12–90
	TMT B-A (s)	36.30 (34.16); range −30–117	49.91 (32.89); range 11–117	15.87 (25.93); range −30–58	22.29 (27.47); range: −46–59
Executive functions	WCST (correct responses)**	39.95 (8.54); range 20–48	36.83 (9.68); range 20–47	44.63 (3.02); range 41–48	45.75 (2.96); range: 38–48
	WCST (categories)	5.95 (1.67); range 3–8	5.33 (1.83); range 3–7	6.88 (0.84); range 6–8	7.19 (0.66); range: 6–8
	WCST (errors)**	8.05 (8.54); range 0–28	11.17 (9.68); range 1–28	3.38 (3.02); range 0–7	2.25 (2.96); range: 0–10
	WCST (perseverations)	2.2 (2.38); range 0–9	3.08 (2.61); range 0–9	.88 (1.13); range 0–3	.75 (1.34); range: 0–4

tailed independent-samples *t*-test (α= 0.05) the study has 80% power to detect between-group differences of approximately Cohen's $d = 0.91$ or larger. For within-group Pearson correlations ($n = 20$), 80% power corresponds to a minimum detectable correlation of $|r| \approx 0.55$. Because these sensitivity estimates are based on parametric assumptions, they represent upper-bound power for the corresponding non-parametric analyses. Moreover, given the application of the Benjamini–Hochberg correction, detectable effects under corrected thresholds are necessarily larger. Accordingly, non-significant findings should be interpreted as reflecting the absence of detectable large effects rather than as evidence for equivalence between groups.

3. Results

3.1. Participants' neuropsychological assessment

The results from the neuropsychological assessment are shown in Table 4. No significant group-related differences were observed in the pretests (MMSE: $U = 190.50$; $p= .317$; $r_{rb}= 0.191$; naming AAT-subtest: $U = 155.50$; $p= .898$; $r_{rb} = -0.028$; token-test: $U = 203.00$; $p= .123$; $r_{rb}= 0.194$) nor in the two tasks assessing working memory (forward digit recall task: $t(34)= 2.032$; $p= .050$; $d = 0.682$; backward digit recall task: $t(34)= -0.085$; $p= .933$; $d= -0.028$). Although the forward digit recall task did not survive correction for multiple comparisons, the associated effect size was moderate, suggesting a possible group-difference that the study may have been underpowered to detect reliably. Thus, no evidence for large group differences emerged in overall basic cognitive status and cognitive-linguistic measures, including working memory, linguistic comprehension, and lexical-access.

However, the assessment of attention and executive functions revealed significant group-related differences. Regarding attention, PwMS were significantly slower than HC in TMT_A ($U = 75.500$; $p < .001$; $r_{rb}= 0.603$). No significant differences were observed between the two groups in TMT_B ($U = 134.000$; $p= .120$; $r_{rb}= 0.295$) and in the TMT_B-A time calculation ($t(37)= -1.407$; $p= .168$; $d= -0.451$). Regarding Executive Functions, PwMS significantly differed from HC in all subtests of the WCST, i.e., correct responses ($U = 240.500$; $p= .010$; $r_{rb} = 0.503$), categories ($U = 231.500$; $p= .015$; $r_{rb} =0.447$), number of errors ($U = 79.500$: $p= .010$; $r_{rb}= -0.503$), and perseverations ($U = 90.000$: $p= .020$; $r_{rb}= -0.438$).

Considering the two subgroups of persons with MS compared to HC, no reliable group differences were observed after correction for multiple comparisons for two pretests (Naming subtest of the AAT: $H(2)= 1.384$, $p= .501$, Rank $\varepsilon^2= 0.040$; Token test: $H(2)= 6.143$, $p= .046$, Rank $\varepsilon^2= 0.171$). Nonetheless, a significant group difference was found in MMSE, $H(2)= 8.880$, $p= .012$, Rank $\varepsilon^2= 0.254$. Dunn's post hoc comparisons revealed that PwMS_High had significantly lower scores than both PwMS_Low ($p= .005$) and HCs ($p= .026$). Considering working memory, the three groups differed on forward digit recall, $H(2)= 10.120$, $p= .006$, Rank $\varepsilon^2= 0.289$, with

Table 5

Results of the Multilevel-Linguistic Analysis of the narrative productions of healthy controls and PwMS and, separately, for the two subgroups of PwMS with high and low MFSI total scores. Results are presented as means (standard deviations). The asterisks show when the group differences were significant after Benjamini–Hochberg correction for multiple comparisons. $p \leq .01$**; $p \leq .001$***. *Legend:* PwMS= Persons with Multiple Sclerosis; HC= Healthy Control participants; PwMS_High= Persons with Multiple Sclerosis and a high level of fatigue; PwMS_Low= Persons with Multiple Sclerosis and a low level of fatigue; MLU= Mean Length of Utterances.

Narrative assessment		PwMS	PwMS_High	PwMS_Low	HC
Productivity	Number of words	118.09 (57.69); range: 39–236	99.01 (45.40); range: 39–173	146.70(65.08); range: 72–236	88.88 (46.67); range: 28–192
	Speech Rate	114.66 (21.15); range: 102–170	105.90(18.10); range: 74–136	127.70(19.31); range: 98–154	129.00 (17.43); range: 102–170
	MLU	6.65 (1.31); range: 5–9	6.00 (1.12); range 5–9	7.58 (1.01); range 6–9	7.48 (1.55); range: 5–12
Lexical Processing	% Phonological Errors***	1.09 (0.73); range: 0–3	1.24 (0.80); range 0–3	.86 (0.57); range 0–2	.31 (0.30); range: 0–1
	% Semantic Errors	1.57 (1.98); range: 0–7	2.32 (2.24); range 0–7	.45 (0.60); range 0–2	.87 (1.47); range: 0–5
Grammatical Processing	% Bound Paragrammatic Errors**	.003 (0.05); range: 0–0.11	.003 (0.003); range: 0–0.11	.00 (0.00); range: 0–0	.001 (0.004); range: 0–0.2
	% Free Paragrammatic Errors	1.08 (1.73); range: 0–8	1.52 (2.13); range 0–8	.41 (0.43); range 0–1	.54 (0.88); range: 0–4
	% Omission of Content Words	17.50 (7.09); range: 6–32	19.97 (7.82); range 6–32	13.79 (3.77); range 7–18	14.92 (7.34); range: 4–30
	% Omission of Function Words	2.28 (1.70); range: 0–6	2.60 (1.79); range 0–6	1.81 (1.55); range 0–4	1.81 (1.94); range: 0–6
	% Grammatical Completeness**	58.49 (11.54); range: 37–84	56.59 (13.73); range 37–84	61.33 (7.05); range 52–70	68.89 (12.67); range: 50–90
Discourse organization	% Cohesion Errors***	23.67 (8.55); range: 8–29	25.22 (10.41); range 8–40	21.35 (4.19); range 16–29	14.89 (6.86); range: 0–29
	% Local Coherence Errors	24.35 (11.25); range: 7–53	28.59 (12.27); range 10–53	17.98 (5.50); range 7–25	20.76 (11.36); range: 8–58
	*% Global Coherence Errors***	15.62 (8.78); range: 3–30	17.87 (8.99); range 4–30	12.24 (7.78); range 3–25	7.83 (5.64); range: 0–18
	% Lexical informativeness**	77.31 (8.21); range: 59–92	75.06 (9.15); range 59–90	80.69 (5.42); range 75–92	85.01 (6.54); range: 73–98

PwMS_High scoring lower than both HC (p= .004) and PwMS_Low (p= .009). However, no group related differences were found for backward digit recall (H(2)= 3.331, p= .189, Rank ε^2= 0.095). As for attention, group-related differences were found only for TMT_A, H(2)= 10.390, p= .006, Rank ε^2= 0.274, with HCs completing the task faster than both PwMS_High (p= .007) and PwMS_Low (p= .010). On TMT_B (H(2)= 6.710, p= .035, Rank ε^2= 0.177) and TMT_B-A (H(2)= 6.736, p= .034, Rank ε^2= 0.177) the three groups performed in a similar way. Finally, regarding executive functions, group-related differences were found also for the number of correct responses at the WCST, H(2)= 9.845, p= .007, Rank ε^2= 0.281, with PwMS_High scoring lower than HCs (p= .005), categories identified at the WCST, H(2)= 8.665, p= .013, Rank ε^2= 0.248, with PwMS_High scoring lower than HCs (p= .003), errors produced at the WCST, H(2)= 9.845, p= .007, Rank ε^2= 0.281, with PwMS_High scoring higher than HCs (p= .002), and perseverative errors produced at the WCST, H(2)= 10.520, p= .005, Rank ε^2= 0.301, with PwMS_High scoring higher than HCs (p= .002) and PwMS_Low (p= .025).

3.2. Narrative assessment through the multilevel linguistic analysis

The results from the narrative assessment are shown in Table 5. The MLA revealed significant group-related differences in several micro- and macro-linguistic variables. Regarding productivity, the two groups did not differ significantly in the number of words (t (38)= - 1.760; p= .086, d= −0.557), speech rate (t(38)= 2.338; p= .025, d = 0.739), or MLU (t(38)= 1.832; p= .075, d = 0.579). This indicates no evidence for large group differences in overall narrative productivity.

As for micro-linguistic variables, at the level of lexical processing the two groups differed significantly in the percentages of phonological errors (U= −68.500; $p < .001$, r_{rb}= -0.658) and of bound paragrammatic errors (U = 102.000; p= .004, r_{rb}= 0.490). In both cases PwMS produced more errors. No significant differences were observed for semantic errors (U = 140.000; p= .098, r_{rb}= 0.300), suggesting no evidence for large impairments in lexical selection skills. Among the total 162 phonological errors (118 by PwMS and 34 by HC), false starts were the most frequent type in both groups (PwMS: n = 112, 69%; HC: n = 34, 21%). Only PwMS produced phonological paraphasias (n = 14, 9%) and neologisms (n = 2, 1%) (see Fig. 2).

At the level of grammatical processing, no significant differences were observed for free paragrammatic errors (U = 131.000; p= .062, r_{rb}= 0.345) or omissions of either content (U = 145.000; p= .142, r_{rb}= 0.275) or function words (U = 161.000; p= .293, r_{rb}= 0.195) supporting the absence of detectable large morphosyntactic difficulties in persons with MS. Nonetheless, the two groups significantly differed in grammatical completeness (t(38)= 2.716; p= .010, d = 0.859), with PwMS producing fewer complete sentences and more errors than HC.

At the level of discourse processing significant group-related differences were observed in cohesion errors (t(38)= -3.585; $p < .001$, d= -1.134), global coherence errors (t-welch (32.4)= +3.336; p= .002, d= −1.055), and lexical informativeness (t(38)= 3.279; p= .002, d = 1.037). PwMS committed more cohesion and global coherence errors and were less appropriate in lexical informativeness compared to HC. No significant differences were found in the production of local coherence errors (U = 148.000; p= .165, r_{rb}= 0.260) supporting the absence of difficulties in establishing local coherence ties between adjacent utterances. Among the 174 total cohesion errors (118 by PwMS and 56 by HC), aposiopesis was the most common type for both groups (PwMS: n = 117, 67%; HC: n = 56, 32%). Only PwMS produced erroneous cohesive functors (n = 1, 1%) (see Fig. 1). Of the 102 global coherence errors observed (65 by PwMS and 37 by HC), filler utterances were the most frequent for both groups (PwMS: n = 39, 38%; HC: n = 22, 21%). Other types of errors included the repetition of propositions (PwMS: n = 24, 24%; HC: n = 14, 14%) and the formulation of conceptually incongruent utterances (PwMS: n = 2; 2%; HC: n = 1, 1%) (see Fig. 2).

Considering the two subgroups of persons with MS, for productivity, the analyses showed that the three groups were no different from each other and performed as HCs on word production (H(2)= 5.872, p= .053, Rank ε^2= 0.151). However, a significant group difference was found in speech rate, H(2)= 10.190, p= .006, Rank ε^2= 0.261. Dunn's post hoc comparisons revealed that PwMS_High had significantly slower speech rates than both PwMS_Low (p= .022) and HCs (p= .002). Similarly, the three groups differed on MLU, H(2)= 11.970, p= .003, Rank ε^2= 0.307, with PwMS_High having significantly lower MLUs than both PwMS_Low (p= .004) and HCs (p= .002).

Group-related differences were found for all three measures assessing lexical processing. For the percentages of phonological errors (H(2)= 13.470, p= .001, Rank ε^2= 0.345) and of semantic errors (H(2)= 7.826, p= .020, Rank ε^2= 0.201), PwMS_High produced more errors than both PwMS_Low (phonological errors: p= .032; semantic errors: p= .025) and HCs (phonological errors: $p < .001$; semantic errors: p= .010). For the percentage of bound paragrammatic errors (H(2)= 9.455, p= .009, Rank ε^2= 0.242) Dunn's post hoc comparisons showed that PwMS_High produced more such errors than HCs (p= .003).

As for grammatical processing, the analyses did not reveal reliable group-related differences after correction for multiple comparisons: free paragrammatic errors, H(2)= 6.066, p= .048, Rank ε^2= 0.156; omissions of content words, H(2)= 4.189, p= .123, Rank ε^2= 0.107; omissions of function words, H(2)= 1.797, p= .407, Rank ε^2= 0.046; complete sentences, H(2)= 5.727, p= .057, Rank ε^2= 0.147.

Finally, considering discourse organization, no differences were found for local coherence errors, H(2)= 5.759, p= .056, Rank ε^2= 0.148. However, group-related differences were found in the percentages of cohesion errors, H(2)= 11.130, p= .004, Rank ε^2= 0.285, with PwMS_High producing significantly more errors than both PwMS_Low (p= .018) and HCs (p= .003), and global coherence errors, H(2)= 8.968, p= .011, Rank ε^2= 0.230, with PwMS_High producing significantly more errors than HCs (p= .003). Finally, a significant group-related differences was found for the percentage of lexical informativeness, H(2)= 9.692, p= .008, Rank ε^2= 0.249, with PwMS_High producing significantly fewer informative words than HCs (p= .003).

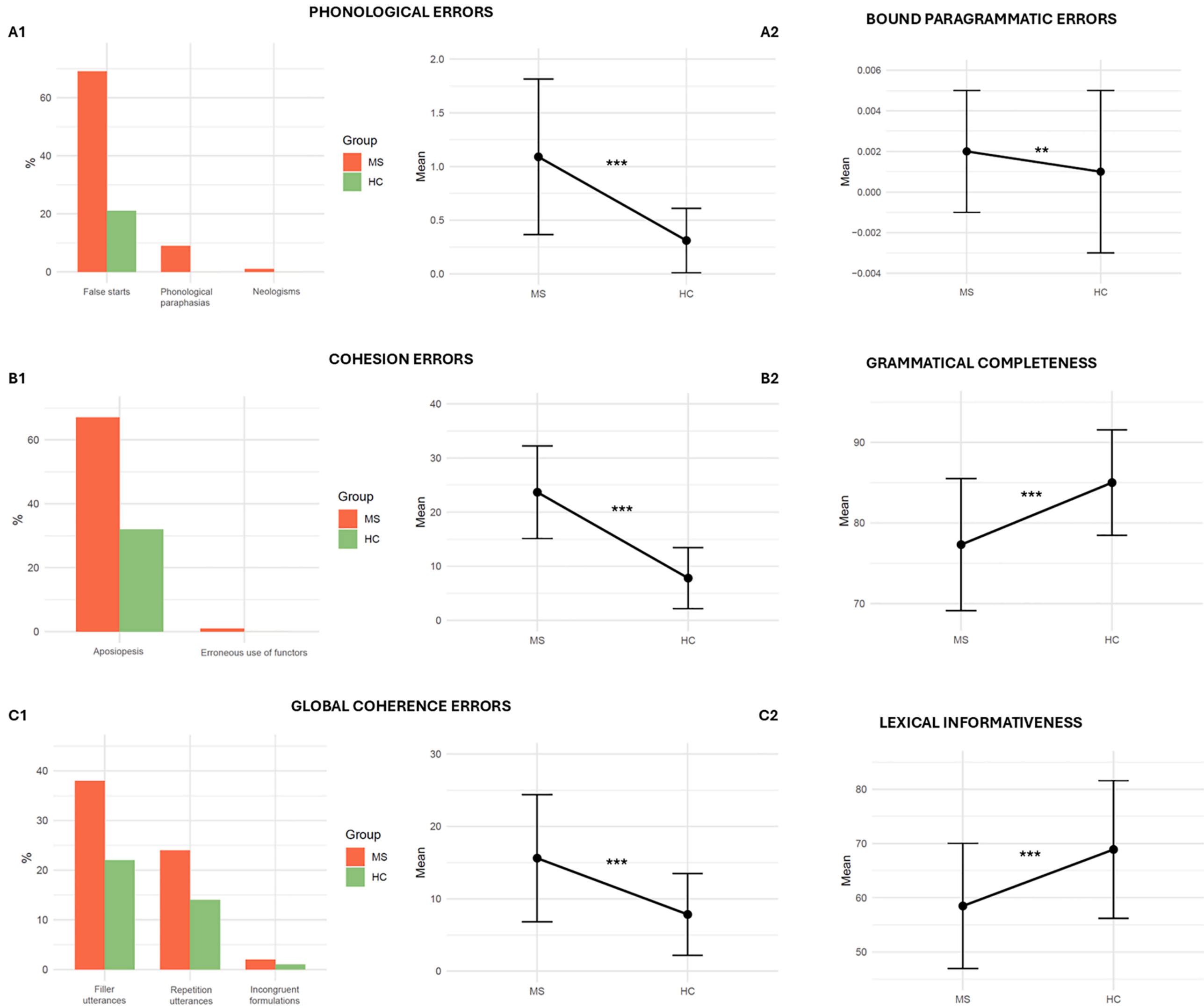

Fig. 2. Representation of the significant group-related differences in micro- and macro-linguistic variables extracted from the MLA. Plots labelled with "1″ (i.e., A1, B1, and C1) show the performance across the two groups in the different types of phonological (i.e., false starts, phonological paraphasias, and neologisms), cohesion (i.e., aposiopesis, and erroneous use of function words), and global coherence errors (i.e., % filler utterances, % repetition utterances, and % incongruent formulations). Plots labelled with "2" (i.e., A2, B2, C2) compare the performance of the two groups on the composite measure of phonological errors, cohesion errors, and global coherence errors, as well as on the % of bound paragrammatic errors, grammatical completeness, and lexical informativeness. **= $p \leq .01$; ***= $p \leq .001$.

3.3. Relationship between cognitive and narrative measures

To explore the potential impact of cognitive impairments on linguistic functioning, a series of correlation analyses was conducted between cognitive measures (i.e., sustained attention and executive functions) and narrative measures (i.e., phonological errors, bound paragrammatic errors, grammatical completeness, cohesion errors, global coherence errors, and lexical informativeness) that were found to be altered in PwMS. No significant correlations were observed between cognitive and linguistic measures in either HC (see Table 6) or PwMS (see Table 7). In light of the sensitivity power analysis, these results indicate the absence of detectable large associations between cognitive and linguistic measures, while smaller or moderate relationships cannot be ruled out.

3.4. Correlation among altered narrative measures in PwMS

Additional correlation analyses were performed among the narrative measures that were found to be altered in PwMS (see Fig. 3). Phonological errors were positively correlated with cohesion errors and bound paragrammatic errors and negatively correlated with grammatical completeness and lexical informativeness. Bound paragrammatic errors were positively correlated with phonological and cohesion errors. Grammatical completeness was negatively correlated with phonological, cohesion and coherence errors and positively correlated with lexical informativeness. Lexical informativeness was negatively correlated with global coherence, cohesion, and phonological errors and positively correlated with grammatical completeness.

4. Discussion

The present study aimed to provide a comprehensive description of the narrative abilities of PwMS. Both PwMS and HC were asked to verbally describe a series of pictures to collect samples of their spontaneous speech. These samples were transcribed and analyzed using an MLA of discourse production, a method proven effective in identifying linguistic difficulties, even in individuals with subtle impairments. Variables assessing micro- and macrolinguistic levels were extracted and compared between the two groups, alongside performance on cognitive tests. Our findings confirmed and expanded previous knowledge about linguistic difficulties in PwMS. Specifically, as hypothesized, we found that (1) the MLA efficiently identified impaired linguistic abilities, even in our cohort of PwMS who showed no difficulties on traditional tasks such as naming and that (2) the MLA demonstrated the interaction between micro- and macrolinguistic difficulties in these participants. Contrary to our hypothesis, (3) no significant relationship was detected between cognitive measures and linguistic difficulties in PwMS, consistent with the absence of detectable large associations in the present sample. Finally, (4) in line with our hypothesis PwMS with higher levels of self-perceived fatigue had reduced efficiency on tasks assessing cognitive and narrative production skills. Importantly, many of the observed linguistic and cognitive alterations were primarily driven by the subgroup of PwMS with higher levels of self-perceived fatigue, whereas PwMS with lower fatigue levels often performed similarly to healthy controls.

Regarding cognitive skills, PwMS showed no evidence of large impairments on tasks assessing working memory but demonstrated difficulties in attention and executive functions. Specifically, while previous studies have reported deficits in both sustained and divided attention among PwMS (McCarthy et al., 2005), participants in this study showed a significant difficulty in sustained attention only (as assessed through the TMT-A). Notably, TMT-A requires sustained attention combined with rapid processing of target stimuli, highlighting processing speed as a critical factor. Unlike TMT-B, TMT-A is less complex, with the primary challenge being the speed of connecting numbers. Thus, the observed impairment in TMT-A may also reflect a general slowdown in processing speed (Litvan et al., 1988), a common cognitive deficit in PwMS (Denney et al., 2005). With respect to executive functions, PwMS exhibited difficulties on all the subtests of the WCST, which specifically assess cognitive flexibility and inhibitory control. This finding aligns with previous research indicating that PwMS tend to make more perseverative errors on this task compared to controls (Parmenter et al., 2007). These results suggest that PwMS may experience general difficulties in abstract reasoning and in utilizing contextual information (i.e., card analysis) to generate goal-oriented behavior (i.e., providing the correct answer).

At the linguistic level, the MLA of discourse production revealed alterations in both micro- and macrolinguistic aspects of language. Notably, our cohort of PwMS was selected to ensure they did not differ from HC in terms of overall cognitive performance, lexical selection and access, or linguistic comprehension. This criterion was essential, as the study aimed to assess the narrative skills of PwMS who did not show linguistic difficulties on standard clinical language assessments, while minimizing the influence of general cognitive impairments (e.g., memory or semantic deficits) on linguistic performance. Interestingly, to the best of our knowledge, this is the first

Table 6
Correlation matrix showing the relationships between cognitive and narrative measures in the group of healthy controls. Spearman's rho and p-values are reported. *Legend*: 1=%Phonological errors; 2=%Grammatical completeness; 3=% Cohesion errors; 4=%Global Coherence Errors; 5=%Lexical informativeness; 6=%Bound paragrammatic errors.

HC	*1*	*2*	*3*	*4*	*5*	*6*
TMT-A (s)	- 0.156; 0.524	.112; 0.646	-0.205; 0.398	.020; 0.935	.084; 0.732	-0.189; 0.437
WCST (correct responses)	.271; 0.309	.128; 0.638	.214; 0.426	-0.495; 0.051	.059; 0.828	-0.258; 0.334
WCST (categories)	.316; 0.233	.142; 0.600	.235; 0.382	-0.393; 0.132	.015; 0.956	-0.174; 0.518
WCST (errors)	-0.271; 0.309	-0.128; 0.638	-0.214; 0.426	.495; 0.051	-0.059; 0.828	.258; 0.334
WCST (perseverations)	-0.033; 0.904	-0.041; 0.88	-0.352; 0.181	.168; 0.535	.222; 0.408	-0.080; 0.768

Table 7
Correlation matrix showing the relationships between cognitive and narrative measures in the group of PwMS. Spearman's rho and p-values are reported. *Legend*: 1 = %Phonological errors; 2 = %Grammatical completeness; 3 = % Cohesion errors; 4 = %Global Coherence Errors; 5 = %Lexical informativeness; 6 = %Bound paragrammatic errors.

PwMS	*1*	*2*	*3*	*4*	*5*	*6*
TMT-A (s)	-0.011; 0.967	.214; 0.364	.023; 0.927	-0.441; 0.053	-0.063; 0.792	.058; 0.808
WCST (correct responses)	-0.131; 0.581	-0.057; 0.813	.044; 0.855	.021; 0.930	-0.033; 0.890	-0.229; 0.330
WCST (categories)	-0.121; 0.611	-0.128; 0.590	.141; 0.554	.125; 0.599	-0.081; 0.735	-0.096; 0.688
WCST (errors)	.131; 0.581	.057; 0.813	-0.044; 0.855	-0.021; 0.930	.033; 0.890	.229; 0.330
WCST (perseverations)	.376; 0.103	-0.184; 0.438	.336; 0.147	-0.038; 0.872	-0.215; 0.363	.386; 0.093

study to document such linguistic difficulties in this population and to highlight the interconnections between such micro- and macrolinguistic impairments.

As for the microlinguistic level, PwMS committed more phonological and bound paragrammatic errors and produced fewer grammatically complete sentences compared to HC. A qualitative analysis revealed that false starts, whose incidence was most pronounced in PwMS with higher fatigue levels, were the most frequent type of phonological errors (see Fig. 2 and Table S1). These errors involve abrupt interruptions in the flow of speech within an ongoing utterance and may be linked to difficulties in other aspects of language production. Indeed, correlation analyses showed that false starts were negatively associated with the percentage of grammatically complete sentences and positively correlated with the production of cohesion errors. A qualitative analysis of cohesion errors revealed that these were primarily aposiopeses (i.e., abrupt interruptions of ongoing utterances). This finding suggests that false starts at the micro-linguistic level are associated with grammatical difficulties and the production of cohesion errors, a macrolinguistic feature. These results align with previous research reporting interactions between micro- and macro-linguistic processes (Andreetta et al., 2012) and underscore the need for accurate, multi-level evaluations of linguistic production in PwMS. Such comprehensive assessments can reveal the intricate relationships between different aspects of language production and provide critical insights for clinical practice.

PwMS produced a slight but significant number of morphological errors, specifically bound paragrammatic errors involving the inappropriate substitution of target bound morphemes. This finding aligns with existing literature on morphological and grammatical difficulties in PwMS. Previous studies reported morphosyntactic difficulties, particularly in verb-tense inflection (Fyndanis et al., 2020; Grigoriadis et al., 2024), suggesting challenges in selecting the appropriate bound morpheme to attach to the infinitive form of verbs. However, in our study, the occurrence of bound paragrammatic errors was minimal, indicating this may represent a trend rather than a clinical marker of the disorder. A possible reason for the discrepancy between our results and those by Fyndanis et al. (2020) and Grigoriadis et al. (2024) lies in the differing methodologies used to elicit morphological processing. In our study, participants were free to produce spontaneous narratives based on picture descriptions, while the other studies used structured tasks explicitly designed to elicit a wide range of morphological structures. Another important distinction may be the grammatical differences between the language studies (Italian in our case and Greek in theirs). While both are inflectional languages, Greek retains case markers that Italian has lost, which may impose different morphological challenges on speakers. This raises an important issue regarding the cross-linguistic generalizability of findings in language studies. Findings from speakers of one language may not directly translate to speakers of another due to differences in grammatical and morphological systems. In this regard, cross-linguistic studies in neurological populations are highly needed to obtain a clear and generalisable description of linguistic difficulties. For example, it is well established that the linguistic symptoms characterizing the non-fluent variant of Primary Progressive Aphasia (nfv-PPA) differ between English and Italian speakers due to language-specific phonological and grammatical differences (Canu et al., 2020; Gorno-Tempini et al., 2011). Specifically, Italian speakers with nfv-PPA tend to produce more morphosyntactic errors in oral production, whereas English speakers with nfv-PPA more frequently exhibit speech-articulatory impairments (Canu et al., 2020). Accordingly, future studies should explicitly consider language-specific characteristics when assessing and interpreting linguistic alterations in PwMS. This cross-linguistic approach will ultimately contribute to a deeper understanding of the psycholinguistic processes affected by the disease.

In our study, the reduced grammatical completeness observed in PwMS was more likely related to difficulties in lexical selection or in accessing semantic information from the mental lexicon, rather than to significant grammatical impairments. This interpretation is supported by the correlation between the production of incomplete sentences and an increased number of false starts, which disrupted sentence flow and hindered the establishment of cohesive ties between utterances. Additionally, grammatical completeness was negatively correlated with cohesion errors (primarily aposiopeses) and global coherence errors, reinforcing the connection between micro- and macrolinguistic difficulties.

Overall, our findings suggest that morphological and syntactic abilities in PwMS were largely preserved at the level of detectable large effects. The observed reduction in grammatical completeness appears to be an epiphenomenon of other challenges, such as increased false starts, which interrupt speech flow and impact discourse cohesion and coherence. This highlights the intricate interplay between micro- and macrolinguistic levels of processing and underscores the importance of assessing these dimensions comprehensively in future studies.

At the macrolinguistic level, PwMS demonstrated reduced discourse cohesion, narrative coherence, and lexical informativeness. Importantly, these effects were most pronounced in PwMS with higher fatigue levels. Among cohesion errors, there were few instances of erroneous use of function words, suggesting that the reduction in cohesion was largely driven by frequent interruptions in their utterances. As noted earlier, these interruptions were associated with reduced grammatical completeness and an increased production

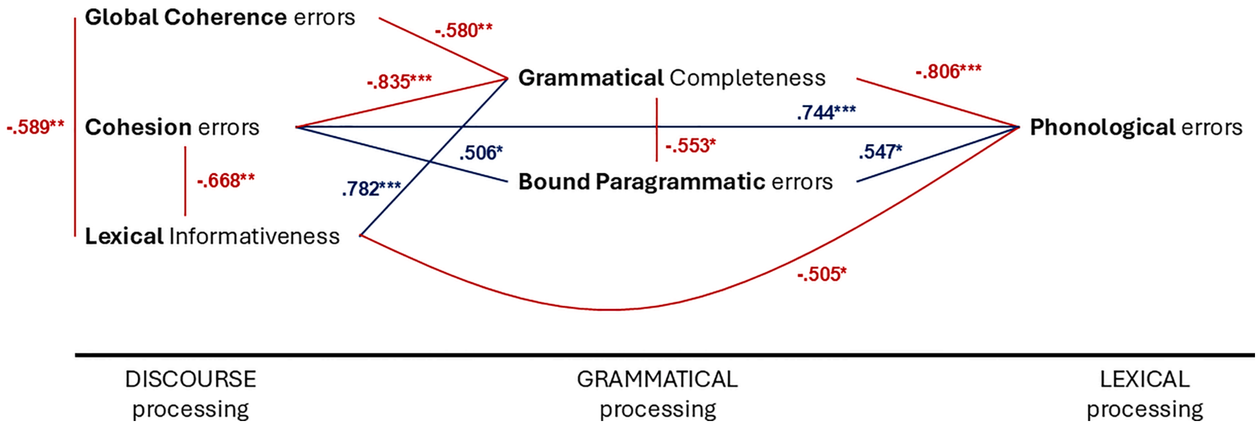

Fig. 3. Spearman rho correlation values. In red negative correlations, in blue positive correlations. $p \leq .05$* $p \leq .01$**; $p \leq .001$***.

of false starts. These issues likely stem from underlying difficulties in discourse planning and the macro-linguistic organization of the narrative flow.

A qualitative analysis of global coherence errors in PwMS revealed that these errors primarily consisted of fillers and repeated utterances with no tangential and few semantically incorrect utterances. Filler and repeated utterances may reflect challenges in narrative organization, as they often indicate an attempt to reformulate or reorient the ongoing story. Global coherence errors were negatively correlated with grammatical completeness and lexical informativeness, suggesting that these errors were more likely to occur in incomplete sentences. It is possible that PwMS produced filler or repeated utterances in sentences that were abruptly interrupted by false starts. This interplay between micro- and macrolinguistic difficulties highlights the impact of sentence-level disruptions on overall narrative structure and coherence. The significant reduction in lexical informativeness observed in PwMS aligns with previous studies reporting that PwMS tend to be under-informative when attempting to convey a message (Arnott et al., 1997; Arrondo et al., 2009; Carotenuto et al., 2018).

When considered from a cross-etiological perspective, the linguistic profile observed in the present study appears to differ from the more pronounced macrolinguistic derailments frequently reported in other populations, such as for persons with moderate-to-severe TBI, where discourse may be marked by tangentiality, severe coherence breakdowns, and reduced goal-directed organization (e.g., Marini et al., 2011). Instead, the present pattern more closely resembles milder discourse alterations described in mild TBI in which fatigue and reduced executive efficiency contribute to subtle cohesion and monitoring deficits without gross narrative disorganization and resulting in the production of filler and repeated utterances. This comparison suggests that, in PwMS, discourse impairments may reflect subtle inefficiencies in monitoring and lexical selection rather than a disruption of macrolinguistic planning mechanisms.

Notably, the present study found a negative correlation between lexical informativeness and global coherence errors. This pattern may reflect an internal strategy employed by PwMS to compensate for difficulties in lexical processing. This hypothesis is supported by a study conducted by Marini and Urgesi (2012), in which non-invasive stimulation (repeated Transcranic Magnetic Stimulation, r-TMS) of the dorsal aspect of the anterior part of the left Inferior Frontal Gyrus (lIFG) in healthy adults temporarily impaired their ability to select appropriate informative words. As a result, participants produced more global coherence errors, potentially as a coping mechanism for perceived lexical difficulties. From this perspective, our findings further validate the self-reported difficulties with word retrieval during spoken discourse (El-Wahsh et al., 2020).

This study also investigated the relationship between impaired cognitive processing and altered narrative measures in PwMS. In light of the sensitivity power analysis, the absence of significant correlations should be interpreted as reflecting the absence of detectable large associations, rather than as evidence against cognitive-linguistic relationships per se. In this regard, heterogeneity within the PwMS group, particularly differences related to fatigue level, may have reduced sensitivity to detect subgroup-specific cognitive-linguistic relationships. Interestingly, no significant correlations were found between measures of sustained attention and executive functions and the micro- and macro-linguistic variables where PwMS performed worse than HC. This finding contrasts with previous studies that reported associations between performance in specific cognitive domains (e.g., social cognition) and linguistic impairments in PwMS (Carotenuto et al., 2018; Delgado-Álvarez et al., 2021). Similarly, studies in other cognitively impaired populations (e.g., Dutta et al., 2024) have indicated that both linguistic and extralinguistic cognitive abilities contribute (depending on the discourse level assessed) to story-retelling performance in individuals with aphasia as well as in the HC group.

Notably, as reported in Tables 6 and Table 7, the present study also found no significant correlations between narrative and cognitive variables in the HC group. This finding adds to the mixed evidence in the existing literature on this topic. Whereas Hilviu et al. (2025), in a sample of healthy older adults, reported significant relationships between linguistic measures, theory of mind abilities and cognitive functions such as working and long-term memory and inhibition, Cannizzaro and Coelho (2013) observed no correlations between narrative measures (i.e., story grammar) and EFs variables, whether linguistically or mixed/not-linguistically based. Nevertheless, the pattern of effect sizes reported by Cannizzaro and Coelho (2013) led the authors to suggest that certain narrative and EF measures may reflect components of a broader construct that similarly influences narrative and executive performance in aging.

From a broader perspective, interpretation of the correlation analyses (particularly in an exploratory study such as the present one) should be approached with caution, given the limited sample size, which may have affected the results. Moreover, a key observation may explain our outcome. The PwMS in our sample were matched with HC in terms of overall cognitive status (i.e., MMSE), naming skills, linguistic comprehension, working memory, and the number of words produced during the story description task. This suggests that in PwMS with mild cognitive symptoms, the relationship between cognitive skills and linguistic performance may parallel that of HC. Moreover, as described above, the PwMS group was distinctly more heterogeneous than the HC group with respect to cognitive variables and other characteristics. This heterogeneity may have influenced the observed relationships between narrative and cognitive measures, suggesting that future studies would benefit from examining these associations in larger and more homogeneous samples. Additionally, an alternative explanation might involve compensatory mechanisms in PwMS that mitigate observable deficits in this population. Finally, another explanation may relate to the type of task employed. Unlike discourse research on traumatic brain injuries, which frequently uses story retell or autobiographical/conversational sampling that taxes working memory, inferencing, and executive control more heavily, in the present study, the picture-description stimulus remained visible throughout the procedure, thereby reducing working memory load relative to story retell or autobiographical narrative paradigms. Under such conditions, associations between executive functions and discourse coherence are less likely to emerge. Accordingly, our paradigm (while well-controlled and sensitive to subtle linguistic alterations) may have underestimated macrolinguistic demands in PwMS with relatively preserved basic cognition. Moreover, differences in neuropathological mechanisms between persons with MS and TBI may further contribute to this divergence. Cognitive deficits in PwMS are often described as diffuse and subcortical, frequently reflecting slowed information processing and distributed white-matter disconnection, whereas persons with TBI commonly involve also frontal damage that more directly disrupts executive monitoring and discourse regulation. Such frontal pathology may yield stronger and more direct associations between executive dysfunction and narrative coherence in persons with TBI. Finally, compensatory mechanisms may operate differently across these etiologies. Given the typically progressive and often slower-evolving course of MS, individuals may develop adaptive strategies that partially mitigate the observable impact of executive inefficiencies on discourse production. In contrast, the abrupt onset of TBI may limit the opportunity for gradual functional reorganization, thereby making executive-discourse relationships more apparent at the behavioral level. Future research should directly investigate hypotheses related to task demands and compensatory mechanisms using more cognitively demanding discourse paradigms and functional neuroimaging techniques.

As noted in the Introduction, the neural underpinnings of spoken discourse remain under investigation (Alyahya, 2023; Schnur et al., 2023). Previous studies have identified a complex and widely distributed neural network involved in spoken discourse, encompassing frontal, anterior and posterior temporal, and parietal regions, along with their underlying white matter tracts (Alyahya, 2023; Braun et al., 2001; Schnur et al., 2023). Consequently, spoken discourse is best understood as a multilevel process requiring contributions from both language-specific networks and non-language regions. When lesions (e.g., stroke) or degeneration (e.g., neurodegenerative or age-related) affect this extensive brain network, neuroplastic compensatory mechanisms may be activated (Barker et al., 2017; Cabeza et al., 2002; Grady, 2012; Hoffman, 2019; Schneider et al., 2021; Wu et al., 2022). For example, Wu and colleagues (2022) found that higher levels of discourse coherence in young adults were predominantly left-lateralized, with increased activation in frontal regions and the superior temporal gyrus during language production. Conversely, Hoffman and colleagues (2019) reported a similar but bilateral activation pattern in healthy older adults. This finding aligns with evidence of compensatory recruitment of the right frontal cortex in healthy aging, compensating for reduced efficiency in the left prefrontal cortex (Cabeza et al., 2002; Grady, 2012). Specifically, highly coherent speech in healthy older adults was associated with increased activity in the bilateral inferior prefrontal cortex (BA45) and rostro-lateral prefrontal cortex (BA10) (Hoffman, 2019). These regions are implicated in selecting lexical information from semantic knowledge (Badre & Wagner, 2007; Gobbo et al., 2021; Thompson-Schill et al., 1997) and planning complex goal-oriented behaviors (Fuster, 2004; Koechlin & Summerfield, 2007). Both cognitive processes showed evidence of vulnerability in our cohort of PwMS, likely contributing to their reduced lexical informativeness and increased global coherence errors.

Previous studies on neurofunctional connectivity in PwMS have reported altered brain activity despite preserved cognitive performance at the behavioral level (Rocca et al., 2022). This altered neurofunctional activity likely reflects compensatory mechanisms that help these persons maintain functional behavior (Helekar et al., 2010). However, such underlying difficulties often become apparent only through demanding tasks, such as the picture description task, combined with a detailed analysis of both micro- and macro-linguistic levels of language processing. The findings of this study underscore the importance of the linguistic production skills of PwMS, even when they do not exhibit linguistic symptoms during traditional assessment tasks. This approach provides a more nuanced perspective, allowing clinicians to detect subtle impairments that may otherwise go unnoticed.

In light of the present findings, it is crucial to expand narrative assessment in both clinical practice and research, as cognitive difficulties may not always align with narrative impairments (and vice versa) as shown in studies on healthy aging (e.g., Bosco et al., 2018; Hilviu et al., 2025. This dissociation has important implications for the design of recovery and enhancement programs, highlighting the need to broaden traditional protocols. Accordingly, speech-language therapy approaches could be refined based on PwMS' needs and the severity of their linguistic alterations. For example, the PwMS enrolled in the present study exhibited linguistic difficulties (namely, in discourse planning, narrative text organization, and lexical-semantic access and selection) that were not detectable through standard tests. Consequently, this group of PwMS may benefit from comprehensive rehabilitative programs that address microlinguistic and macrolinguistic levels (see for example Parola et al. 2019) in an integrated perspective, thereby exploiting the inherent complexity of discourse production (Dipper et al., 2021; Linnik et al., 2016). More specifically, these findings suggest that PwMS with higher fatigue levels may particularly benefit from discourse-based interventions targeting narrative planning, cohesion, and lexical selection. Rather than focusing narrowly on a single linguistic aspect (e.g., lexical access through naming tasks), intervention programs could incorporate activities that simultaneously recruit micro- and macrolinguistic processes. For example, conversational settings may be used to stimulate the production of complex syntactic structures and the retrieval of low-frequency words by encouraging discussion of selected topics (e.g., Gobbo et al., 2024). Finally, because individuals with mild linguistic difficulties are often still actively engaged in work and social activities (Mozeiko & Pascariello, 2020), therapeutic interventions should target higher-level cognitive-linguistic processes to meet the communicative demands of professional environments and complex discourse contexts (Laks et al., 2025).

Despite its relevance and its novelty, our study is not without limitations. First, the relatively small sample size warrants caution and highlights the need for future research involving larger cohorts. Additionally, although recruitment challenges are well recognized, multicentric studies would enhance the generalizability of the current findings and allow for more fine-grained analyses of cognitive and narrative variables in subgroups with more homogeneous characteristics (e.g., a more detailed categorization of fatigue levels).

In conclusion, this study provided a comprehensive assessment of the narrative abilities of PwMS, identifying both strengths and weaknesses that warrant attention. Using an MLA of discourse production, the study detailed micro- and macro-linguistic difficulties in PwMS, demonstrating that their linguistic impairments span lexical, grammatical, and discourse-level processes. These findings underscore the importance of incorporating training programs that target both micro- and macro- linguistic components of narrative production into clinical care paths.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Marika Gobbo: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Ilaria Gabbatore:** Writing – review & editing, Writing – original draft, Conceptualization. **Giulia Piovani:** Methodology, Investigation, Data curation. **Francesca Marina Bosco:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Andrea Marini:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Data curation, Conceptualization. **Andrea Tacchino:** Writing – review & editing, Supervision, Methodology, Investigation, Data curation, Conceptualization.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrea Marini reports financial support was provided by Italian Ministry of Research. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by PRIN 2022 PNRR, Prot. n. P2022M9JCM, project title: "Standardization of the Multilevel prOcedure for discOurse analysis and Training program for narrative production in Healthy adults - SMOOTH". Avvisopubblico n. 1409 del 14/09/2022 – PRIN 2022 PNRR M4C2 Inv. 1.1. Ministero dell'Università e della Ricerca (Financed by EU, NextGenerationEU) – CUP G53D23007250001

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jcomdis.2026.106657.

References

Adornetti, I., Chiera, A., Altavilla, D., Deriu, V., Marini, A., Gobbo, M., Valeri, G., Magni, R., & Ferretti, F. (2023). Defining the characteristics of story production of autistic children: A multilevel analysis. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-023-06096-2

Aj, T., Se, B., J, G., B, H., & O, C. (2018). Multiple sclerosis. *Lancet (London, England), 391*(10130). https://doi.org/10.1016/S0140-6736(18)30481-1

Alyahya, R. S. W. (2023). The structural neural correlates of spoken discourse. In A. P.-H. Kong (Ed.), *Spoken discourse impairments in the neurogenic populations: A state-of-the-art, contemporary approach* (pp. 111–119). Springer International Publishing. https://doi.org/10.1007/978-3-031-45190-4_8.

Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia, 50*(8), 1787–1793. https://doi.org/10.1016/j.neuropsychologia.2012.04.003

Andreetta, S., Marini, A., Menichelli, A., Furlanis, G., Vincis, E., Caruso, P., Naccarato, M., & Manganotti, P. (2024). Language assessment in persons with aphasia early after thrombolysis: The utility of multilevel procedures of discourse analysis. *Aphasiology*. https://doi.org/10.1080/02687038.2024.2340797

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*. https://doi.org/10.1076/1380-3395(200008)22:4;1-0;FT518

Arcara, G., & Bambini, V. (2016). A test for the assessment of pragmatic abilities and cognitive substrates (APACS): Normative data and psychometric properties. *Frontiers in Psychology, 7*, 70. https://doi.org/10.3389/fpsyg.2016.00070

Arnett, P. A., & Strober, L. B. (2011). Cognitive and neurobehavioral features in multiple sclerosis. *Expert Review of Neurotherapeutics, 11*(3), 411–424. https://doi.org/10.1586/ern.11.12

Arnott, W. L., Jordan, F. M., Murdoch, B. E., & Lethlean, J. B. (1997). Narrative discourse in multiple sclerosis: An investigation of conceptual structure. *Aphasiology, 11*(10), 969–991. https://doi.org/10.1080/02687039708249422

Arrondo, G., Sepulcre, J., Duque, B., Toledo, J., & Villoslada, P. (2009). Narrative speech is impaired in multiple sclerosis. *European Neurological Journal*.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883–2901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Bara, B. G. (2010). *Cognitive pragmatics*. MIT Press.

Barker, M. S., Young, B., & Robinson, G. A. (2017). Cohesive and coherent connected speech deficits in mild stroke. *Brain and Language, 168*, 23–36. https://doi.org/10.1016/j.bandl.2017.01.004

Beatty, W. W., Wilbanks, S. L., Blanco, C. R., Hames, K. A., Tivis, R., & Paul, R. H. (1996). Memory disturbance in multiple sclerosis: Reconsideration of patterns of performance on the selective reminding test. *Journal of Clinical and Experimental Neuropsychology, 18*(1), 56–62. https://doi.org/10.1080/01688639608408262

Bosco, F. M., Tirassa, M., & Gabbatore, I. (2018). Why pragmatics and theory of mind do not /(completely) overlap. *Frontiers in psychology, 9*, 1453.

Boucher, J., Brisebois, A., Slegers, A., et al. (2022). Picture description of the Western Aphasia battery picnic scene: Reference data for the French Canadian population. *American Journal of Speech-Language Pathology, 31*(1), 257–270. https://doi.org/10.1044/2021_AJSLP-20-00388

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample-size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37*(2), 399–407. https://doi.org/10.1044/jshr.3702.399

Braun, A. R., Guillemin, A., Hosey, L., & Varga, M. (2001). The neural organization of discourse: An H2 15O-PET study of narrative production in english and American sign language. *Brain: A Journal of Neurology, 124*(Pt 10), 2028–2044. https://doi.org/10.1093/brain/124.10.2028

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics, 30*(7), 489–518. https://doi.org/10.3109/02699206.2016.1145740

Büttner-Kunert, J., Blöchinger, S., Falkowska, Z., Rieger, T., & Oslmeier, C. (2022). Interaction of discourse processing impairments, communicative participation, and verbal executive functions in people with chronic traumatic brain injury. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.892216

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage, 17*(3), 1394–1402. https://doi.org/10.1006/nimg.2002.1280

Cannizzaro, M. S., & Coelho, C. A. (2013). Analysis of narrative discourse structure as an ecologically relevant measure of executive function in adults. *Journal of Psycholinguistic Research, 42*(6), 527–549. https://doi.org/10.1007/s10936-012-9231-5

Canu, E., Agosta, F., Battistella, G., Spinelli, E. G., DeLeon, J., Welch, A. E., Mandelli, M. L., Hubbard, H. I., Moro, A., Magnani, G., Cappa, S. F., Miller, B. L., Filippi, M., & Gorno-Tempini, M. L. (2020). Speech production differences in english and italian speakers with nonfluent variant PPA. *Neurology, 94*(10), e1062–e1072. https://doi.org/10.1212/WNL.0000000000008879

Carletta, J. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics, 22*(2), 249–254. http://dl.am.org/citation.cfm?id=230386.230390.

Carotenuto, A., Arcara, G., Orefice, G., Cerillo, I., Giannino, V., Rasulo, M., Iodice, R., & Bambini, V. (2018). Communication in multiple sclerosis: Pragmatic deficit and its relation with cognition and social cognition. *Archives of Clinical Neuropsychology, 33*(2), 194–205. https://doi.org/10.1093/arclin/acx061

Chiaravalloti, N. D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet. Neurology, 7*(12), 1139–1151. https://doi.org/10.1016/S1474-4422(08)70259-X

David, W. (1987). *Manual for the wechsler memory scale-revised*. Psychological Corporation.

De Renzi, E., & Vignolo, L. A. (1962). The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain: A Journal of Neurology, 85*, 665–678. https://doi.org/10.1093/brain/85.4.665

Delgado-Álvarez, A., Matias-Guiu, J. A., Delgado-Alonso, C., Hernández-Lorenzo, L., Cortés-Martínez, A., Vidorreta, L., Montero-Escribano, P., Pytel, V., & Matias-Guiu, J. (2021). Cognitive processes underlying verbal fluency in multiple sclerosis. *Frontiers in Neurology, 11*. https://doi.org/10.3389/fneur.2020.629183

DeLuca, J., Chelune, G. J., Tulsky, D. S., Lengenfelder, J., & Chiaravalloti, N. D. (2004). Is speed of processing or working memory the primary information processing deficit in multiple sclerosis? *Journal of Clinical and Experimental Neuropsychology, 26*(4), 550–562. https://doi.org/10.1080/13803390490496641

Denney, D. R., Sworowski, L. A., & Lynch, S. G. (2005). Cognitive impairment in three subtypes of multiple sclerosis. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 20*(8), 967–981. https://doi.org/10.1016/j.acn.2005.04.012

Diamond, B. J., DeLuca, J., Kim, H., & Kelley, S. M. (1997). The question of disproportionate impairments in visual and auditory information processing in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology, 19*(1), 34–42. https://doi.org/10.1080/01688639708403834

Dipper, L., Marshall, J., Boyle, M., Botting, N., Hersh, D., Pritchard, M., & Cruice, M. (2021). Treatment for improving discourse in aphasia: A systematic review and synthesis of the evidence base. *Aphasiology, 35*(9), 1125–1167. https://doi.org/10.1080/02687038.2020.1765305

Dutta, M., Murray, L. L., & Stark, B. C. (2024). The relationship between executive functioning and narrative language abilities in aphasia. *American Journal of Speech-Language Pathology*, 1–24. https://doi.org/10.1044/2024_AJSLP-23-00314

El-Wahsh, S., Heard, R., & Bogaardt, H. (2019). Development and validation of a speech pathology-specific questionnaire for persons with multiple sclerosis (SMS). *International Journal of Speech-Language Pathology, 21*(6), 553–563. https://doi.org/10.1080/17549507.2018.1499802

El-Wahsh, S., Ballard, K., Kumfor, F., & Bogaardt, H. (2020). Prevalence of self-reported language impairment in multiple sclerosis and the association with health-related quality of life: An international survey study. *Multiple Sclerosis and Related Disorders, 39*, Article 101896. https://doi.org/10.1016/j.msard.2019.101896

El-Wahsh, S., Heard, R., Bogaardt, H., Kumfor, F., & Ballard, K. J. (2021). Variables associated with self-reported language impairment in multiple sclerosis: A regression analysis. *International Journal of MS Care, 23*(2), 85–92. https://doi.org/10.7224/1537-2073.2020-096

Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., & Schlech, W. F. (1994). Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, 18*(Suppl 1), S79–S83. https://doi.org/10.1093/clinids/18.supplement_1.s79

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Foong, J., Rozewicz, L., Quaghebeur, G., Davie, C. A., Kartsounis, L. D., Thompson, A. J., Miller, D. H., & Ron, M. A. (1997). Executive function in multiple sclerosis. The role of frontal lobe pathology. *Brain: A Journal of Neurology, 120*(Pt 1), 15–26. https://doi.org/10.1093/brain/120.1.15

Fuster, J. M. (2004). Upper processing stages of the perception-action cycle. *Trends in Cognitive Sciences, 8*(4), 143–145. https://doi.org/10.1016/j.tics.2004.02.004

Fyndanis, V., Messinis, L., Nasios, G., Dardiotis, E., Martzoukou, M., Pitopoulou, M., Ntoskou, A., & Malefaki, S. (2020). Impaired verb-related morphosyntactic production in multiple sclerosis: Evidence from Greek. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.02051

Giovagnoli, A. R., Del Pesce, M., Mascheroni, S., Simoncelli, M., Laiacona, M., & Capitani, E. (1996). Trail making test: Normative values from 287 normal adult controls. *The Italian Journal of Neurological Sciences, 17*(4), 305–309. https://doi.org/10.1007/BF01997792

Gobbo, M., Agostini, L., Priftis, K., & De Pellegrin, S. (2024). A clinical diary and timeline to monitor joint therapy for apraxia of speech and aphasia. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*. https://doi.org/10.1007/s10072-024-07758-4

Gobbo, M., De Pellegrin, S., Bonaudo, C., Semenza, C., Della Puppa, A., & Salillas, E. (2021). Two dissociable semantic mechanisms predict naming errors and their responsive brain sites in awake surgery. DO80 revisited. *Neuropsychologia, 151*, Article 107727. https://doi.org/10.1016/j.neuropsychologia.2020.107727

Goodglass, H., & Kaplan, E. (1972). *The boston diagnostic aphasia examination*. Lea & Febiger.

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience, 13*(7), 491–505. https://doi.org/10.1038/nrn3256

Grigoriadis, P., Bakirtzis, C., Nteli, E., Boziki, M.-K., Kotoumpa, M., Theotokis, P., Kesidou, E., & Stavrakaki, S. (2024). Morphosyntactic abilities and cognitive performance in multiple sclerosis. *Brain Sciences, 14*(3). https://doi.org/10.3390/brainsci14030237

Helekar, S. A., Shin, J., Mattson, B. J., Bartley, K., Stosic, M., Saldana-King, T., Montague, R., & Hutton, G. J. (2010). Functional brain network changes associated with maintenance of cognitive function in Multiple sclerosis. *Frontiers in Human Neuroscience, 4*. https://doi.org/10.3389/fnhum.2010.00219

Hilviu, D., Parola, A., Bosco, F. M., Marini, A., & Gabbatore, I. (2025). Grandpa, tell me a story! narrative ability in healthy aging and its relationship with cognitive functions and Theory of Mind. *Language, Cognition and Neuroscience, 40*(1), 103–121. https://doi.org/10.1080/23273798.2024.2401027

Hoffman, P. (2019). Reductions in prefrontal activation predict off-topic utterances during speech production. *Nature Communications, 10*, 515. https://doi.org/10.1038/s41467-019-08519-0

Huber, W., & Gleber, J. (1982). Linguistic and nonlinguistic processing of narratives in aphasia. *Brain and Language, 16*(1), 1–18. https://doi.org/10.1016/0093-934X(82)90069-4

Kertesz, A. (1982). *The western aphasia battery*. Grune & Stratton.

Kobelt, G., Berg, J., Atherly, D., & Hadjimichael, O. (2006). Costs and quality of life in multiple sclerosis: A cross-sectional study in the United States. *Neurology, 66*(11), 1696–1702. https://doi.org/10.1212/01.wnl.0000218309.01322.5c

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11*(6), 229–235. https://doi.org/10.1016/j.tics.2007.04.005

Kong, A., Cheung, P.-H., C, Y.-N., & Wong, C. W-Y (2025). Establishing norm of connected speech measures for descriptive discourses in Cantonese-speaking adults. *International Journal of Language & Communication Disorders, 60*, Article e70055. https://doi.org/10.1111/1460-6984.70055

Kristensson, J., Longoni, F., Östberg, P., Åke, S., Rödseth Smith, S., & Saldert, C. (2022). Evaluation of treatment effects of semantic feature analysis on mild anomia in multiple sclerosis. *Aphasiology, 36*(6), 706–731. https://doi.org/10.1080/02687038.2021.1900535

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology, 33*(11), 1444–1452. https://doi.org/10.1212/wnl.33.11.1444

Labov, W., & Waletsky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and the visual arts* (pp. 3–38). Seattle and London: University of Washington Press.

Laks, A. B., DeMarco, A. T., & Turkeltaub, P. E. (2025). Optimizing detection of very mild aphasia with letter fluency tests. *American Journal of Speech-Language Pathology, 34*(3), 1193–1202. https://doi.org/10.1044/2025_AJSLP-24-00471

Linnik, A., Bastiaanse, R., & Höhle, B. (2016). Discourse production in aphasia: A current review of theoretical and methodological challenges. *Aphasiology, 30*(7), 765–800. https://doi.org/10.1080/02687038.2015.1113489

Larson, R. D. (2013). Psychometric properties of the modified fatigue impact scale. *Ineternational Journal of MS Care, 15*(1), 15–20. https://doi.org/10.7224/1537-2073.2012-019

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Litvan, I., Grafman, J., Vendrell, P., & Martinez, J. M. (1988). Slowed information processing in multiple sclerosis. *Archives of Neurology, 45*(3), 281–285. https://doi.org/10.1001/archneur.1988.00520270059021

Luzzatti, C., Poeck, K., Weniger, D., Huber (psicologo), W., Bleser, R. D., & Willmes, K. (1996). *AAT: Aachener aphasie test: Manuale e dati normativi*. Organizzazioni speciali.

Marcotte, K., Roy, a., Brisebois, A., Jutras, C., Leonard, C., Rochon, E., & Brambati, S. (2024). Reliability of the picture description task of the Western Aphasia Battery – revised in laurentian French persons without brain injury. *The Clinical Neuropsychologist, 38*(8), 1980–2008. https://doi.org/10.1080/13854046.2024.2340777

Marini, A. (2012). Characteristics of narrative discourse processing after damage to the right hemisphere. *Seminars in Speech and Language, 33*(1), 68–78. https://doi.org/10.1055/s-0031-1301164

Marini, A. (2022). Cognitive and linguistic characteristics of narrative discourse production in healthy aging. In C. Coelho, L. R. Cherney, & B. Shadden (Eds.), *Discourse analysis in adults with and without communication disorders: A resource for clinicians and researchers* (pp. 15–31). Plural Publishing, Inc.

Marini, A. (2023). Experimental methods to study disorders of language production in adults. In S. Zufferey, & Gygax (Eds.), *The Routledge handbook of experimental linguistics* (pp. 407–421). Routledge. https://doi.org/10.4324/9781003392972-30.

Marini, A., Andreetta, S., del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology, 25*(11), 1372–1392. https://doi.org/10.1080/02687038.2011.584690

Marini, A., Galetto, V., Zampieri, E., Vorano, L., Zettin, M., & Carlomagno, S. (2011). Narrative language in traumatic brain injury. *Neuropsychologia, 49*(10), 2904–2910. https://doi.org/10.1016/j.neuropsychologia.2011.06.017

Marini, A., Zettin, M., Bencich, E., Bosco, F. M., & Galetto, V. (2017). Severity effects on discourse production after TBI. *Journal of Neurolinguistics, 44*, 91–106. https://doi.org/10.1016/j.jneuroling.2017.03.005

Marini, A., Petriglia, F., D'Ortenzio, S., Bosco, F. M., & Gasparotto, G. (2025). Unveiling the dynamics of discourse production in healthy aging and its connection to cognitive skills. *Discourse Processes, 62*(6–7), 479–501. https://doi.org/10.1080/0163853X.2025.2507548

Marini, A., Petriglia, F., Gasparotto, G., D'Ortenzio, S., Andreetta, S., & Gobbo, M. (2026). Standardization of the multiLevel discourse Analysis (MLA) and identification of data-driven age bands for research on language in healthy aging. *Journal of Neurolinguistics, 79*, 101343. https://doi.org/10.1016/j.neuroling.2026.101343

Marini, A., Spoletini, I., Rubino, I. A., Ciuffa, M., Bria, P., Martinotti, G., Banfi, G., Boccascino, R., Strom, P., Siracusano, A., Caltagirone, C., & Spalletta, G. (2008). The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research, 105*(1), 144–155. https://doi.org/10.1016/j.schres.2008.07.011

Marini, A., & Urgesi, C. (2012). Please get to the point! A cortical correlate of linguistic informativeness. *Journal of Cognitive Neuroscience, 24*(11), 2211–2222. https://doi.org/10.1162/jocn_a_00283

Mazzon, G., Ajčević, M., Cattaruzza, T., Menichelli, A., Guerriero, M., Capitanio, S., Pesavento, V., Dore, F., Sorbi, S., Manganotti, P., & Marini, A. (2019). Connected speech deficit as an early hallmark of CSF-defined Alzheimer's disease and correlation with cerebral hypoperfusion pattern. *Current Alzheimer Research, 16*(6), 483–494. https://doi.org/10.2174/1567205016666190506141733

McCarthy, M., Beaumont, J. G., Thompson, R., & Peacock, S. (2005). Modality-specific aspects of sustained and divided attentional performance in multiple sclerosis. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 20*(6), 705–718. https://doi.org/10.1016/j.acn.2005.04.007

Merlini, A., Gabbatore, I., Marini, A., & Bosco, F. M. (2026). Affective and cognitive theory of mind in aging: What's the role of executive functions and cognitive reserve on the decline? *Experimental Aging Research, 52*(3), 377–405. https://doi.org/10.1080/0361073X.2025.2586970

Migliore, S., Curcio, G., Couyoumdjian, A., Ghazaryan, A., Landi, D., Moffa, F., Quintiliani, L., Squitieri, F., Palmieri, M. G., Filippi, M. M., & Vernieri, F. (2018). Executive functioning in relapsing-remitting multiple sclerosis patients without cognitive impairment: A task-switching protocol. *Multiple Sclerosis (Houndmills, Basingstoke, England), 24*(10), 1328–1336. https://doi.org/10.1177/1352458517719149

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. *Neurological Sciences, 34*(5), 749–754. https://doi.org/10.1007/s10072-012-1130-x

Mozeiko, J., & Pascariello, A. (2020). How are SLPs managing services for people with mild aphasia? *Journal of Communication Disorders, 84*, Article 105983. https://doi.org/10.1016/j.jcomdis.2020.105983

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior, 12*(4), 313–324. https://doi.org/10.1016/s0010-9452(76)80035-4

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech, Language, and Hearing Research, 36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338

Parmenter, B. A., Zivadinov, R., Kerenyi, L., Gavett, R., Weinstock-Guttman, B., Dwyer, M. G., Garg, N., Munschauer, F., & Benedict, R. H. B. (2007). Validity of the Wisconsin card sorting and Delis-Kaplan Executive function system (DKEFS) sorting tests in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology, 29*(2), 215–223. https://doi.org/10.1080/13803390600672163

Parola, A., Bosco, F. M., Gabbatore, I., Galetto, V., Zettin, M., & Marini, A. (2019). The impact of the Cognitive Pragmatic Treatment on the pragmatic and informative skills of individuals with traumatic brain injury (TBI). *Journal of Neurolinguistics, 51*, 53–62.

Perlini, C., Marini, A., Garzitto, M., Isola, M., Cerruti, S., Marinelli, V., Rambaldelli, G., Ferro, A., Tomelleri, L., Dusi, N., Bellani, M., Tansella, M., Fabbro, F., & Brambilla, P. (2012). Linguistic production and syntactic comprehension in Schizophrenia and Bipolar Disorder. *Acta Psychiatrica Scandinavica, 126*(5), 363–376. https://doi.org/10.1111/j.1600-0447.2012.01864.x

Plotas, P., Nanousi, V., Kantanis, A., Tsiamaki, E., Papadopoulos, A., Tsapara, A., Glyka, A., Mani, E., Roumelioti, F., Strataki, G., Fragkou, G., Mavreli, K., Ziouli, N., & Trimmis, N. (2023). Speech deficits in multiple sclerosis: A narrative review of the existing literature. *European Journal of Medical Research, 28*(1). https://doi.org/10.1186/s40001-023-01230-3

Portaccio, E., & Amato, M. P. (2022). Cognitive impairment in multiple sclerosis: An update on assessment and management. *NeuroSci, 3*(4), 667–676. https://doi.org/10.3390/neurosci3040048

Rao, S. M. (1995). Neuropsychology of multiple sclerosis. *Current Opinion in Neurology, 8*(3), 216.

Renauld, S., Mohamed-Saïd, L., & Macoir, J. (2016). Language disorders in multiple sclerosis: A systematic review. *Multiple Sclerosis and Related Disorders, 10*, 103–111. https://doi.org/10.1016/j.msard.2016.09.005

Rocca, M. A., Schoonheim, M. M., Valsasina, P., Geurts, J. J. G., & Filippi, M. (2022). Task- and resting-state fMRI studies in multiple sclerosis: From regions to systems and time-varying analysis. Current status and future perspective. *Neuroimage-Clinical, 35*, Article 103076. https://doi.org/10.1016/j.nicl.2022.103076

Schneider, F., Marcotte, K., Brisebois, A., Townsend, S. A. M., Smidarle, A. D., Loureiro, F., da Rosa Franco, A., Soder, R. B., Nikolaev, A., Marrone, L. C. P., & Hübner, L. C. (2021). Neuroanatomical correlates of macrolinguistic aspects in narrative discourse in unilateral left and right hemisphere stroke: A voxel-based morphometry study. *Journal of Speech, Language, and Hearing Research: JSLHR, 64*(5), 1650–1665. https://doi.org/10.1044/2020_JSLHR-20-00500

Schnur, T. T., Brown, D. A., & Guess, K. K. (2023). The functional neural correlates of spoken discourse. In A. P.-H. Kong (Ed.), *Spoken discourse impairments in the neurogenic populations: A state-of-the-art, contemporary approach* (pp. 121–132). Springer International Publishing. https://doi.org/10.1007/978-3-031-45190-4_9.

Schuell, H., Jenkins, J. J., & Jimenez-Pabon, E. (1964). *Aphasia in adults: Diagnosis, prognosis, and treatment*. New York, NY: Hoeber Medical Division, Harper & Row.

Spalletta, G., Spoletini, I., Cherubini, A., Rubino, I. A., Siracusano, A., Piras, F., Caltagirone, C., & Marini, A. (2010). Cortico-subcortical underpinnings of narrative processing impairment in schizophrenia. *Psychiatry Research, 182*(1), 77–80. https://doi.org/10.1016/j.pscychresns.2009.11.001

Swirsky-Sacchetti, T., Mitchell, D. R., Seward, J., Gonzales, C., Lublin, F., Knobler, R., & Field, H. L. (1992). Neuropsychological and structural brain lesions in multiple sclerosis. *Neurology, 42*(7), 1291. https://doi.org/10.1212/WNL.42.7.1291. –1291.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792

Wu, W., Morales, M., Patel, T., Pickering, M. J., & Hoffman, P. (2022). Modulation of brain activity by psycholinguistic information during naturalistic speech comprehension and production. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior, 155*, 287–306. https://doi.org/10.1016/j.cortex.2022.08.002

Yorkston, K. M., Baylor, C. R., Klasner, E. R., Deitz, J., Dudgeon, B. J., Eadie, T., Miller, R. M., & Amtmann, D. (2007). Satisfaction with communicative participation as defined by adults with multiple sclerosis: A qualitative study. *Journal of Communication Disorders, 40*(6), 433–451. https://doi.org/10.1016/j.jcomdis.2006.10.006